



TUGAS AKHIR - SS141501

**PREDIKSI PEMBELOTAN KONSUMEN SOFTWARE ANTIVIRUS
'X' DENGAN BINARY LOGISTIC REGRESSION DAN LOGISTIC
REGRESSION ENSEMBLES**

Ayu Asfihani
NRP 1311 100 084

Dosen Pembimbing
Dr.rer.pol. Heri Kuswanto, M.Si

Program Studi S1 Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember
Surabaya 2015



FINAL PROJECT - SS141501

**PREDICTION OF CUSTOMER DEFECTION OF ANTIVIRUS 'X'
WITH BINARY LOGISTIC REGRESSION AND LOGISTIC
REGRESSION ENSEMBLES**

Ayu Asfihani
NRP 1311 100 084

Supervisor
Dr.rer.pol. Heri Kuswanto, M.Si

Undergraduate Programme of Statistics
Faculty of Mathematics and Natural Science
Institut Teknologi Sepuluh Nopember
Surabaya 2015

LEMBAR PENGESAHAN

PREDIKSI PEMBELOTAN KONSUMEN SOFTWARE ANTIVIRUS 'X' DENGAN BINARY LOGISTIC REGRESSION DAN LOGISTIC REGRESSION ENSEMBLES

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada

Program Studi S-1 Jurusan Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh :

AYU ASFIHANI
NRP. 1311 100 084

Disetujui oleh Pembimbing Tugas Akhir :

Dr.rer.pol. Heri Kuswanto, M.Si

NIP: 19810326 200312 1 004

Mengetahui :

Ketua Jurusan Statistika FMIPA-ITS

Dr. Muhammad Mashuri, M.T

NIP: 19620408 198701 1 001

SURABAYA, JULI 2015

PREDIKSI PEMBELOTAN KONSUMEN SOFTWARE ANTIVIRUS 'X' DENGAN BINARY LOGISTIC REGRESSION DAN LOGISTIC REGRESSION ENSEMBLES

Nama Mahasiswa : Ayu Asfihani
NRP : 1311 100 084
Jurusan : Statistika FMIPA-ITS
Dosen Pembimbing : Dr.rer.pol. Heri
Kuswanto, M.Si

ABSTRAK

Penelitian pembelotan konsumen Perusahaan 'X' sangat dibutuhkan karena mengukur loyalitas konsumen tidaklah mudah, hal ini terutama disebabkan oleh Big Data dan kelas pembelotan konsumen yang tidak seimbang. Setelah sebelumnya diteliti oleh Prasasti, dkk (2013), Prasasti dan Ohwada (2014), dan Martono dkk (2014) dengan Machine Learning, pembelotan konsumen Perusahaan 'X' diteliti dengan Binary Logistic Regression dan LORENS untuk membentuk model hingga melakukan pengklasifikasian. Dengan 500000 konsumen Low Price, 408810 konsumen Medium Price, dan 709899 konsumen High Price, hubungan antara variabel Akumulasi Update, Harga Produk, Jawaban Kontrak, Tipe Konsumen, dan Status Pengiriman dengan Pembelotan Konsumen dijelaskan oleh Binary Logistic Regression. Namun metode tersebut tidak mampu menangani Big Data dalam tahap pengujian parameter. Harga Produk dan Tipe Konsumen terbukti berpengaruh signifikan terhadap model namun tidak dapat menjelaskan kecenderungan pembelotan konsumen Medium Price dan High Price. Karena itu, hasil analisis dan klasifikasi dengan Binary Logistic Regression dalam penelitian ini kurang dapat dipercaya. Dengan kemampuan menangani Big Data, ketidak-seimbangan respon, serta ketimpangan jumlah prediktor terhadap banyak pengamatan, LORENS digunakan. Meskipun begitu, LORENS tidak mampu menjelaskan hubungan antar variabel karena tidak dapat menghasilkan model yang interpretatif. Meskipun akurasi klasifikasi Binary Logistic Regression terlihat sedikit lebih baik di konsumen Low Price dan Medium Price daripada LORENS (Low Price 66,54 % : 66,25%, Medium Price 77,32% : 74,06 %, High Price 68,42% : 69,04%), metode yang disarankan untuk diaplikasikan untuk data ini adalah LORENS karena lebih dapat dipercaya. Cross Validation yang disarankan LORENS hanya mampu meningkatkan performa LORENS pada konsumen High Price dengan akurasi 69,21%.

Kata kunci : *Big Data, Binary Logistic Regression, Cross Validation, Holdout, Klasifikasi, LORENS*

Halaman ini sengaja dikosongkan.

PREDICTION OF CUSTOMER DEFECTION OF ANTIVIRUS 'X' SOFTWARE WITH BINARY LOGISTIC REGRESSION AND LOGISTIC REGRESSION ENSEMBLES

Name : Ayu Asfihani
NRP : 1311 100 084
Department : Statistika FMIPA-ITS
Supervisor : Dr.rer.pol. Heri
Kuswanto, M.Si

ABSTRACT

This research about customer's defection of 'X' Company is really needed because measuring their loyalties is not easy, especially because of Big Data and imbalanced customer's defection responses. After being researched by Prasasti, dkk (2013), Prasasti dan Ohwada (2014), and Martono dkk (2014) with Machine Learning, in this research it will be treated with Binary Logistic Regression and LORENS to form model until perform classification. With 5000 00 Low Price customers, 408810 Medium Price customers, and 709899 High Price customers, relation between variables of Akumulasi Update, Harga Produk, Jawaban Kontrak, Tipe Konsumen, and Status Pengiriman to Pembelotan Konsumen is explained by Binary Logistic Regression. But this method can't handle Big Data in parameter's significance testing. Harga Produk and Tipe Konsumen are proven significantly effecting to model but can't explain any propensity of Medium Price and High Price customer's defection. Thus, analysis and classification result from Binary Logistic Regression in this research is less trustworthy. With the ability to handle Big Data, imbalanced responses, and unfairness of predictor's number to observation's amounts, LORENS is applied. Nonetheless, LORENS can't explain any relation between variables because it doesn't produce any interpretative model. Although classification accuracy from Binary Logistic Regression seems little bit better in Low Price and Medium Price customers than in LORENS (Low Price 66,54 % : 66, 25%, Medium Price 77, 32% : 74, 06%, High Price 68,42% : 69,04%), the suggested method to be applied in this case is LORENS because it is more trustworthy. Cross Validation, as suggested evaluation method by LORENS only able to enhance LORENS's performance in High Price customers with accuracy 69,21%.

Key Words : *Big Data, Binary Logistic Regression, Classification, Cross Validation, Holdout, LORENS*

Halaman ini sengaja dikosongkan.

KATA PENGANTAR

Bismillahirrohmanirrohim,

Puji syukur penulis panjatkan ke hadirat Allah SWT atas rahmat dan hidayah-Nya penulis dapat menyelesaikan Tugas Akhir yang berjudul **“Prediksi Pembelotan Konsumen Software Antivirus ‘X’ Dengan Binary Logistic Regression Dan Logistic Regression Ensembles”**. Tugas akhir ini disusun sebagai salah satu syarat memperoleh gelar kesarjanaan S-1 jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember.

Terselesaikannya Tugas Akhir ini tidak terlepas dari peran serta dari berbagai pihak. Oleh karena itu penulis mengucapkan terima kasih kepada :

1. Dr.rer.pol. Heri Kuswanto, M.Si selaku dosen pembimbing utama yang bersedia meluangkan waktu untuk memberikan bimbingan, masukan, dan arahan dalam menyelesaikan Tugas Akhir ini.
2. Dr. Muhammad Mashuri, M.T selaku ketua Jurusan Statistika ITS yang telah mendukung dan membimbing sehingga Tugas Akhir ini dapat terselesaikan.
3. Dra. Sri Mumpuni Retnaningsih, M.T selaku dosen wali yang banyak memberikan dukungan serta kemudahan kepada Penulis selama menjalani studi.
4. Dra. Wiwiek Setya Winahju. M.S dan Dr. Irhamah. S.Si. M.Si selaku dosen penguji yang telah memberikan saran dan kritik sehingga mendorong penulis untuk selalu belajar.
5. Seluruh dosen Jurusan Statistika ITS yang telah memberikan banyak ilmu serta segenap karyawan Jurusan Statistika ITS yang mendukung terselesaikannya Tugas Akhir ini.

6. Pihak-pihak lain yang telah mendukung dan membantu atas terseselaikannya Tugas Akhir ini yang tidak bisa disebutkan satu persatu.

Semoga laporan Tugas Akhir ini bermanfaat dan menambah wawasan keilmuan kita semua. Penulis mengharapkan saran dan kritik yang membangun untuk perbaikan dimasa yang akan datang demi kebaikan bersama.

Surabaya, Juli 2015

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	ii
TITLE PAGE	iii
LEMBAR PENGESAHAN.....	iv
ABSTRAK.....	v
ABSTRACT.....	vii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xi
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xv
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian.....	6
1.4 Batasan Masalah.....	6
BAB II TINJAUAN PUSTAKA.....	7
2.1 <i>Binary Logistic Regression (BLR)</i>	7
2.2 <i>Logistic Regression Classification by Ensembles from Random Partitions (LR CERP)</i>	12
2.3 <i>Logistic Regression Ensembles (LORENS)</i>	14
2.4 <i>Holdout dan Cross validation</i>	18
2.5 <i>Pembelotan Konsumen Perusahaan ‘X’</i>	21
2.6 <i>Penelitian Sebelumnya</i>	23
BAB III METODOLOGI PENELITIAN	27
3.1 <i>Sumber Data</i>	27
3.2 <i>Variabel Penelitian</i>	25
3.3 <i>Langkah Analisis</i>	29
BAB IV ANALISIS DAN PEMBAHASAN.....	35

4.1 Analisis Statistika Deskriptif	35
4.2 Analisis <i>Binary Logistic Regression</i>	40
4.3 Analisis LORENS	47
4.4 Analisis LORENS dengan <i>Cross Validation</i>	56
4.4 Pemilihan Metode Terbaik	62
BAB V KESIMPULAN DAN SARAN	65
5.1 Kesimpulan	65
5.2 Saran	67
DAFTAR PUSTAKA	69
LAMPIRAN.....	71
BIODATA PENULIS	95

DAFTAR TABEL

	Halaman
Tabel 2.1	Tabulasi Silang Klasifikasi Prediksi dan Klasifikasi Sebenarnya 16
Tabel 2.2	Variabel Asli dari Situs <i>E-commerce</i> Perusahaan ‘X’ 22
Tabel 2.3	Perbedaan Akurasi dari Penelitian Sebelumnya.. 24
Tabel 3.1	Struktur Data Penelitian 28
Tabel 4.1.1	Statistik Deskriptif Variabel Harga Produk 35
Tabel 4.1.2	Statistik Deskriptif Variabel Akumulasi Update 36
Tabel 4.2.1	Koefisien Parameter Model Awal untuk Konsumen <i>Low Price</i> 40
Tabel 4.2.2	Koefisien Parameter Model Terbaik untuk Konsumen <i>Low Price</i> 41
Tabel 4.2.3	<i>Odss Ratio</i> Koefisien Parameter Model Terbaik untuk Konsumen <i>Low Price</i> 41
Tabel 4.2.4	Koefisien Parameter Model Terbaik untuk Konsumen <i>Medium Price</i> 42
Tabel 4.2.5	<i>Odss Ratio</i> Koefisien Parameter Model Terbaik BLR untuk Data <i>Medium Price</i> 42
Tabel 4.2.6	Koefisien Parameter Model Terbaik untuk data <i>High Price</i> 44
Tabel 4.2.7	<i>Odss Ratio</i> Koefisien Parameter Model Terbaik untuk Data <i>High Price</i> 44
Tabel 4.2.8	Tabulasi Silang Klasifikasi dari <i>Binary Logistic Regression</i> 46
Tabel 4.2.9	Ketepatan Klasifikasi Analisis <i>Binary Logistic Regression</i> 46
Tabel 4.2.10	Kesalahan Klasifikasi Analisis <i>Binary Logistic</i>

	<i>Regression</i>	46
Tabel 4.3.1	<i>Random Sampling</i> Variabel Prediktor pada Ruang Partisi	48
Tabel 4.3.2	Model Regresi Logistik pada Ruang Partisi	49
Tabel 4.3.3	Tabulasi Silang Klasifikasi Konsumen <i>Low Price</i> untuk 3 Partisi dan <i>Threshold</i> Optimum.....	51
Tabel 4.3.4	Tabulasi Silang Klasifikasi Hasil Analisis LORENS untuk Konsumen <i>Low Price</i>	52
Tabel 4.3.5	Tabulasi Silang Klasifikasi Hasil Analisis LORENS untuk Konsumen <i>Medium Price</i>	52
Tabel 4.3.6	Tabulasi Silang Klasifikasi Hasil Analisis LORENS untuk Konsumen <i>High Price</i>	53
Tabel 4.3.7	Ukuran Ketepatan Klasifikasi Hasil Analisis LORENS	54
Tabel 4.3.8	Ukuran Partisi Optimum dari Analisis LORENS	55
Tabel 4.3.9	Ukuran Kesalahan Klasifikasi Hasil Analisis LORENS	56
Tabel 4.4.1	<i>Threshold</i> Optimum Konsumen <i>High Price</i> untuk 2 Partisi.....	57
Tabel 4.4.2	Tabulasi Silang Klasifikasi Konsumen <i>High Price</i> dengan <i>Cross Validation</i> untuk 2 Partisi dan <i>Threshold</i> Optimum.....	58
Tabel 4.4.3	Tabulasi Silang Klasifikasi Hasil Analisis LORENS dengan <i>Cross Validation</i> untuk Konsumen <i>Low Price</i>	59
Tabel 4.4.4	Tabulasi Silang Klasifikasi Hasil Analisis LORENS dengan <i>Cross Validation</i> untuk Konsumen <i>Medium Price</i>	59

Tabel 4.4.5	Tabulasi Silang Klasifikasi Hasil Analisis LORENS dengan <i>Cross Validation</i> untuk Konsumen <i>High Price</i>	60
Tabel 4.4.6	Ukuran Ketepatan Klasifikasi Hasil Analisis LORENS dengan <i>Cross Validation</i>	60
Tabel 4.4.7	Ukuran Kesalahan Klasifikasi Hasil Analisis LORENS dengan <i>Cross Validation</i>	61
Tabel 4.4.8	Ukuran Partisi Optimum dari Analisis LORENS dengan <i>Cross Validation</i>	62
Tabel 4.5.1	Perbandingan Ketepatan Klasifikasi	62

Halaman ini sengaja dikosongkan.

DAFTAR GAMBAR

	Halaman
Gambar 2.1	Bagan Konsep LR CERP 13
Gambar 2.2	Diagram Alir LORENS..... 17
Gambar 3.1	Diagram Alir Analisis Penelitian 31
Gambar 3.2	Diagram Alir Analisis Penelitian (Lanjutan Pertama)..... 32
Gambar 3.3	Diagram Alir Analisis Penelitian (Lanjutan Kedua) 33
Gambar 4.1	Diagram Batang Jawaban Kontrak..... 37
Gambar 4.2	Diagram Batang Tipe Konsumen 38
Gambar 4.3	Diagram Batang Status Pengiriman 39
Gambar 4.4	Diagram Batang Pembelotan Konsumen..... 39

Halaman ini sengaja dikosongkan.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sejak ditemukan pada tahun 1969, internet telah berkembang pesat dan telah terlibat dalam sebagian besar hidup manusia, mulai dari urusan pribadi hingga urusan kelompok sekalipun. Keberadaan internet yang reliabel, fleksibel, dan dapat diakses dengan mudah, membawa manusia ke keadaan yang ramai akan umpan (*feed*). Jam kerja internet dalam menghasilkan umpan secara *real time* adalah salah satu sebab terbentuknya *Big Data*, suatu kondisi dimana data yang dikumpulkan kaya akan variabel dengan jumlah pengamatan yang tidak terbatas. Hal tersebut terjadi pada semua *Cloud Based Software* (perangkat lunak berbasis internet) yang tengah memamerkan beberapa tren yang sangat nyata. 36% pertumbuhan tahunan *cloud software* di pasar perangkat lunak diperkirakan akan terus berlanjut sampai 2016 (Colombus, 2013). Peningkatan pertumbuhan *software* ini didukung oleh kenyamanan *software* yang artinya dapat digunakan di mana-mana selama perangkat pengguna terhubung ke internet.

Dengan meningkatnya penggunaan internet dalam beberapa tahun terakhir ini, urusan memprediksi keputusan konsumen tengah sangat diperhatikan oleh banyak peneliti dan sangat penting bagi perusahaan terutama bagi yang menggunakan sistem kontrak dengan konsumennya. Penelitian selama ini difokuskan pada mencari metode yang paling efektif memprediksi keputusan pelanggan. Metode yang paling umum digunakan untuk memprediksi adalah *Decision Tree*, Regresi, *Naive Bayes*, dan *Neural Network*. Perusahaan 'X' adalah organisasi yang menjual produk antivirus yang beroperasi dengan sambungan internet (*Cloud Based Software*). Perusahaan 'X' bekerja dengan sistem kontrak dan tengah menyelesaikan masalah konsumen yang membelot. Dalam kasus ini, membelot diartikan tidak lagi menggunakan suatu produk antivirus apapun dari perusahaan 'X' yang ditunjukkan dari pemberhentian atau perpanjangan kontrak. Seperti yang dialami

oleh banyak perusahaan dengan data yang disediakan dengan kemudahan akses dari internet, *Big Data* juga tengah dihadapi oleh Perusahaan 'X'.

Di penelitian sebelumnya Martono, Kanamori, dan Ohwada (2014) menyebutkan bahwa tugas memprediksi konsumen Perusahaan 'X' bukanlah hal yang mudah karena empat alasan berikut. Pertama, fitur data yang terbatas dan mencakup hanya beberapa atribut konsumen. Kedua, hampir tidak mungkin untuk mendapatkan informasi konsumen lainnya dengan langsung mendekati setiap konsumen, karena perusahaan ini memiliki banyak konsumen. Ketiga, data yang tersedia hanya mengandung catatan aktivitas konsumen dalam memilih *in* (melanjutkan) dan memilih *out* (meninggalkan) dari suatu produk, sementara pada kenyataannya beberapa konsumen memilih *out* untuk meng-*upgrade* /*downgrade* produk yang mereka pakai. Keempat, dengan pertumbuhan pasar yang luas, mengelola pembelotan pelanggan adalah isu penting. Dalam penelitian ini, perilaku konsumen akan dipelajari dari beberapa karakteristik seperti frekuensi *update* yang dilakukan, harga produk yang dibeli, jenis konsumen, status pengiriman *e-mail* dan status kontrak konsumen. Karakteristik tersebut diwakili oleh fitur Akumulasi Update, Harga Produk, Jawaban Kontrak, Tipe Konsumen, Status Pengiriman yang berperan sebagai variabel prediktor dan Pembelotan Konsumen yang berperan sebagai variabel respon.

Masalah dasar memprediksi pembelotan pelanggan adalah menemukan model yang baik yang dapat memprediksi dengan tepat. Metode statistika dapat menangani masalah prediksi klasifikasi dan pendekatan komputasi dapat meringankan penanganan kasus *Big Data*.

Big Data telah menjadi tantangan bagi para analis. Pendekatan parametrik yang menggunakan uji signifikansi dan statistik inferensia yang disertai dengan asumsi-asumsi yang harus dipenuhi sering kali menjadi lemah bila dihadapkan pada *Big Data*. Lin, Lucas, dan Shmueli (2013) telah menulis tentang *P-Value* yang sensitif terhadap banyaknya pengamatan. Statistika infe-

rensia didasarkan pada gagasan hipotesa *null* yang umumnya menunjukkan bahwa variabel tertentu tidak memberikan pengaruh yang signifikan. Tantangan data yang besar hadir saat pembuktian bahwa hipotesa *null* dapat ditolak. Limitasi nilai pengamatan menuju tak hingga mengakibatkan standard error menuju 0 dan statistik uji menuju tak hingga. Uji dengan data yang besar cenderung menyatakan parameter yang diuji berpengaruh signifikan karena *P-Value* cenderung jatuh menuju 0. Dengan sampel yang besar, beberapa pendekatan parametrik minimal menjadi tidak efektif bahkan tidak berguna, namun kasus yang paling dihindari adalah kesimpulan analisis yang menyesatkan. Metode pengolahan *BigData* semakin berkembang ke pendekatan komputasional yang umumnya tidak menggunakan pengujian dan tidak membuktikan asumsi. Dengan mengadaptasi ide pendekatan inferensia yang lemah dan menyempurnakannya dengan algoritma yang agregatif, seringkali kondisi pendekatan komputasional lebih dapat dilakukan dan hasilnya lebih dapat diandalkan.

Prasasti, Okada, Kanamori, dan Ohwada (2013) memprediksi pembelotan konsumen Perusahaan 'X' dengan teknik *Machine Learning*. Metode klasifikasi yang digunakan adalah *C4.5 Decision Tree* dan *Support Vector Machine (SVM)* yang keduanya memberikan akurasi tertinggi pada produk *HighPrice*. Pada tahun 2014 Prasasti dan Ohwada meneliti kembali kasus ini dengan metode klasifikasi *J48 Decision Tree (J48)*, *Random Forest (RF)*, *Neural Network* dengan fungsi *Multi LayerPerception (MLP)*, dan *Support Vector Machine* dengan algoritma *SMO*. *J48*, *RF* dan *SMO* menghasilkan akurasi terbaik pada produk *High Price*, sedangkan *MLP* memberikan hasil terbaik pada produk *Medium Price*. Pada tahun yang sama Martono dkk (2014) melanjutkan penelitian ini dengan metode klasifikasi *C4.5 Decision Tree* yang menghasilkan akurasi terbaik pada produk *High Price*, penelitian inilah yang menghasilkan akurasi terbaik pada produk *Low Price* dan *High Price*. Sedangkan akurasi terbaik produk *Medium Price* dihasilkan oleh *Random Forest*.

Suatu metode klasifikasi baru telah dikembangkan dari metode klasifikasi statistika dasar *Logistic Regression* (LR) untuk dua kelas yang membolehkan variabel kategori termasuk sebagai prediktornya. Metode baru ini dikembangkan oleh Lim pada tahun 2007 dengan menyertakan algoritma *Classification by Ensembles from Random Partition* (CERP) yang mempartisi variabel-variabel penentu keputusan menjadi beberapa subruang kemudian menggabungkan kembali model-model berbasis *Logistic Regression* (LR) dari masing-masing partisi ke dalam satu fungsi. Metode baru tersebut bernama *Logistic Regression Ensembles* (LORENS).

Hal lain yang dihadapi Perusahaan 'X' adalah banyak pengamatan konsumen pada ketiga jenis produk berupa membelot dan tidak membelot yang tidak seimbang. King dan Zeng (2001) menjelaskan bahwa kelas respon tidak seimbang mengakibatkan estimasi parameter dalam *Binary Logistic Regression* yang menggunakan *Maximum Likelihood Estimator* menjadi bias karena probabilitas kelas tidak membelot akan membesar seiring kecilnya probabilitas konsumen membelot. Dalam hal ini matriks *Hessian* yang digunakan dalam estimasi akan semakin kecil dan hasil estimasi menjadi bias. Selain itu, *threshold* 0,5 yang digunakan dalam klasifikasi menjadi tidak adil terhadap probabilitas masing-masing kelas. LORENS menyediakan *threshold* optimum yang adil terhadap kelas respon pengamatan.

Keunggulan LORENS dibangun dari sifat LR yang informatif dan representatif serta sifat CERP yang membuat variabel penentu keputusan menjadi saling *mutually exclusive* (Lee, Ahn, Moon, Kodell, dan Chen, 2013). Penelitian LORENS sebelumnya pernah dilakukan oleh Lim (2010) kepada data penderita *Acute Mycloid Leukemia* (AML). Dari penelitian tersebut dijelaskan bahwa LORENS dapat dibandingkan dengan metode klasifikasi lain yang telah disebutkan di atas dalam hal meningkatkan akurasi, *sensitivity*, dan *specificity*. Meskipun LORENS lahir untuk menangani kondisi data berdimensi tinggi dimana banyak prediktor lebih besar, penelitian ini dilakukan untuk menguji kesesuaian

metode LORENS terhadap data Perusahaan ‘X’ yang banyak pengamatannya lebih besar dari banyak prediktornya. Hasil prediksi LORENS akan dibandingkan dengan hasil klasifikasi metode *Binary Logistic Regression* yang merupakan versi regresi logistik untuk kasus klasifikasi yang terdiri atas dua kelas dan merupakan basis klasifikasi yang diadaptasi oleh metode LORENS. Metode-metode komputasi seperti LORENS memang mampu memperbaiki hasil klasifikasi, tetapi tidak memberikan informasi yang lengkap tentang model yang dibangun. *Binary Logistic Regression* penting dilakukan dalam penelitian ini karena interpretasi yang menjelaskan hubungan antara variabel-variabel prediktor dengan pembelotan konsumen butuh untuk dijelaskan.

1.2 Rumusan Masalah

Pembelotan konsumen telah menjadi perhatian Perusahaan ‘X’ semenjak konsumen-konsumen yang membelot tidak dapat diprediksi secara langsung. Setelah diteliti oleh Prasasti, dkk (2013), Prasasti dan Ohwada (2014), dan Martono dkk (2014) dengan *Machine Learning*, kasus ini akan diteliti dengan *Binary Logistic Regression* dan LORENS. Dengan respon biner, masalah dalam penelitian ini adalah mengklasifikasikan konsumen Perusahaan antivirus ‘X’ menjadi membelot dan tidak membelot. Namun *Binary Logistic Regression* sebagai pendekatan inferensial dapat memberikan hasil yang kurang dapat dipercaya dalam menangani kasus *Big Data* karena cenderung menolak hipotesa *null* pada tahap pengujian parameter. Maka dalam mengklasifikasikan konsumen Perusahaan ‘X’, metode LORENS sebagai pendekatan komputasional yang tidak menggunakan uji parameter akan digunakan. Namun metode LORENS tidak memberikan *out-put* model yang dapat menggambarkan hubungan antar variabel prediktor dengan variabel respon. Maka *Binary Logistic Regression* digunakan dalam rangka mendapatkan model regresi dan hasil klasifikasi. Pemilihan metode akan ditujukan pada metode yang baik dalam analisis dan mampu menghasilkan ketepatan klasifikasi pembelotan konsumen dengan baik.

1.3 Tujuan Penelitian

Berdasarkan rumusan permasalahan yang dipaparkan di atas, berikut merupakan tujuan dari penelitian ini.

1. Mendapatkan model *Binary Logistic Regression* untuk menggambarkan hubungan antara variabel Akumulasi Update, Harga Produk, Jawaban Kontrak, Tipe Konsumen, dan Status Pengiriman dengan variabel respon yang berupa Pembelian Konsumen. Serta mendapatkan hasil klasifikasi dan tingkat ketepatan klasifikasi perilaku konsumen antivirus 'X' dengan metode tersebut.
2. Mendapatkan hasil klasifikasi dan tingkat ketepatan klasifikasi perilaku konsumen antivirus 'X' menggunakan LOR-ENS.
3. Memilih metode terbaik di antara hasil analisis dari *Binary Logistic Regression* dan LORENS.

1.4 Batasan Masalah

Hal yang membatasi masalah dalam penelitian ini adalah *preprocessing* Data (*Analisis Missing Value* dan *Variable Selection*) telah dilakukan di penelitian sebelumnya.

BAB II TINJAUAN PUSTAKA

2.1 *Binary Logistic Regression*

Regresi logistik merupakan metode klasifikasi dasar yang awalnya diperuntukkan kepada variabel respon dengan dua kelas yang bernama *Binary Logistic Regression*, kemudian berkembang untuk kasus dengan variabel respon yang terdiri atas multi kelas yang bernama *Multinomial Logistic Regression*. Regresi logistik atau regresi logistik biner merupakan suatu metode analisis data yang digunakan untuk mencari hubungan antara variabel respon (y) yang bersifat *biner* atau dikotomis dengan variabel prediktor (x) yang bersifat polikotomis (Hosmer dan Lemeshow, 2000). Variabel respon y terdiri dari 2 kategori yaitu “sukses” dan “gagal” yang dinotasikan dengan $y = 1$ (sukses) dan $y = 0$ (gagal). Sehingga variabel y mengikuti distribusi Bernoulli untuk setiap observasi tunggal. Fungsi probabilitas untuk setiap observasi diberikan sebagai berikut.

$$f(y) = \pi^y (1 - \pi)^{1-y}; y = 0, 1 \quad (2.1)$$

dimana jika $y = 0$ maka $f(y) = 1 - \pi$ dan jika $y = 1$ maka $f(y) = \pi$. Fungsi regresi logistik dapat dituliskan sebagai berikut.

$$f(z) = \frac{e^{z=\beta_0+\beta_1x_1+\dots+\beta_px_p}}{1 + e^{z=\beta_0+\beta_1x_1+\dots+\beta_px_p}} \quad (2.2)$$

dimana p adalah banyaknya variabel prediktor. Nilai z terletak antara $-\infty$ dan $+\infty$ sehingga nilai $f(z)$ terletak antara 0 dan 1 untuk setiap nilai z yang diberikan. Hal tersebut menunjukkan bahwa model logistik sebenarnya menggambarkan probabilitas atau resiko dari suatu objek. Model regresi logistiknya adalah sebagai berikut.

$$\pi(x) = \frac{e^{(\beta_0+\beta_1x_1+\dots+\beta_px_p)}}{1 + e^{(\beta_0+\beta_1x_1+\dots+\beta_px_p)}} \quad (2.3)$$

Untuk memudahkan pendugaan parameter regresi maka model regresi logistik dapat diuraikan dengan menggunakan transformasi logit dari persamaan ini $\{\pi(x)\} \{1 + e^{(\beta_0+\beta_1x_1+\dots+\beta_px_p)}\} = e^{(\beta_0+\beta_1x_1+\dots+\beta_px_p)}$ sehingga dapat diperoleh persamaan berikut.

$$g(x) = \ln \left(\frac{\pi(x)}{1 - \pi(x)} \right) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (2.4)$$

Dalam model regresi linier, diasumsikan bahwa amatan dari variabel respon diekspresikan sebagai $y = E(Y|x) + \varepsilon$ dimana $E(Y|x) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$ merupakan rata-rata dari populasi dan ε merupakan penyimpangan amatan dari rata-ratanya dan diasumsikan berdistribusi $N(0, \sigma^2)$. Namun pada regresi logistik variabel respon diekspresikan sebagai $y = \pi(x) + \varepsilon$ dimana ε mempunyai salah satu dari kemungkinan dua nilai yaitu $\varepsilon = 1 - \pi(x)$ dengan peluang $\pi(x)$ jika $y = 1$ dan $\varepsilon = -\pi(x)$ dengan peluang $1 - \pi(x)$ jika $y = 0$ yang mengikuti distribusi binomial dengan rata-rata nol dan varians $(\pi(x))(1 - \pi(x))$.

Estimasi parameter dalam regresi logistik menggunakan *Maximum Likelihood*. metode tersebut mengestimasi parameter β dengan cara memaksimumkan fungsi *likelihood* dan mensyaratkan bahwa data harus mengikuti suatu distribusi tertentu yang dalam regresi logistik adalah Bernoulli.

Jika x_i dan y_i adalah pasangan variabel bebas dan terikat pada pengamatan ke- i dan diasumsikan bahwa setiap pengamatan saling independen dengan pasangan pengamatan lainnya, maka fungsi probabilitas untuk setiap pasangan adalah sebagai berikut.

$$f(x_i) = \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i}; y_i = 0, 1 \quad (2.5)$$

dengan,

$$\pi(x_i) = \frac{e^{(\sum_{j=0}^p \beta_j x_j)}}{1 + e^{(\sum_{j=0}^p \beta_j x_j)}} \quad (2.6)$$

dimana bila $j = 0$ maka nilai $x_{i0} = x_{i0} = 1$.

Setiap pasangan pengamatan diasumsikan independen sehingga fungsi *likelihood* merupakan gabungan dari fungsi distribusi masing-masing pasangan sebagai berikut.

$$\begin{aligned} l(\beta) &= \prod_{i=1}^n f(x_i) = \prod_{i=1}^n \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \\ &= \left\{ \prod_{i=1}^n (1 - \pi(x_i)) \right\} \left\{ \prod_{i=1}^n e^{\left(\log \left(\frac{\pi(x_i)}{1 - \pi(x_i)} \right) \right)^{y_i}} \right\} \end{aligned}$$

$$\begin{aligned}
&= \left\{ \prod_{i=1}^n (1 - \pi(x_i)) \right\} e^{\left\{ \sum_{i=1}^n y_i \log \left(\frac{\pi(x_i)}{1 - \pi(x_i)} \right)^{y_i} \right\}} \\
&= \left\{ \prod_{i=1}^n \frac{1}{1 + e^{\sum_{j=0}^p \beta_j x_{ij}}} \right\} e^{\left\{ \sum_{i=1}^n y_i \log \left(e^{\sum_{j=0}^p \beta_j x_{ij}} \right) \right\}} \\
&= \left\{ \prod_{i=1}^n \left(1 + e^{\sum_{j=0}^p \beta_j x_{ij}} \right)^{-1} \right\} e^{\left\{ \sum_{j=0}^p \left(\sum_{i=1}^n y_i x_{ij} \right) \beta_j \right\}} \quad (2.7)
\end{aligned}$$

Fungsi *likelihood* tersebut lebih mudah dimaksimumkan dalam bentuk $\log l(\beta)$ dan dinyatakan dengan $L(\beta)$.

$$L(\beta) = \log l(\beta) = \sum_{j=0}^p \left(\sum_{i=1}^n y_i x_{ij} \right) \beta_j - \sum_{i=1}^n \log \left(1 + e^{\sum_{j=0}^p \beta_j x_{ij}} \right) \quad (2.8)$$

Nilai β maksimum didapatkan melalui turunan $L(\beta)$ terhadap β dan hasilnya adalah sama dengan 0.

$$\begin{aligned}
\frac{\partial L(\beta)}{\partial \beta_j} &= \sum_{i=1}^n y_i x_{ij} - \sum_{i=1}^n x_{ij} \left(\frac{e^{\sum_{j=0}^p \beta_j x_{ij}}}{1 + e^{\sum_{j=0}^p \beta_j x_{ij}}} \right) \\
&= \sum_{i=1}^n y_i x_{ij} - \sum_{i=1}^n x_{ij} \hat{\pi}(x_i) = 0; \quad j = 0, 1, \dots, p \quad (2.9)
\end{aligned}$$

Estimasi varians dan kovarians dikembangkan melalui teori MLE dari koefisien parameter (Rao, 1973 dalam Hosmer dan Lemeshow, 1989). Teori tersebut menyatakan bahwa estimasi varians kovarians didapatkan melalui turunan kedua $L(\beta)$.

$$\frac{\partial L(\beta)}{\partial \beta_j \partial \beta_u} = \sum_{i=1}^n x_{ij} x_{iu} \pi(x_i) (1 - \pi(x_i)); \quad j, u = 0, 1, \dots, p \quad (2.10)$$

Matriks varians kovarians berdasarkan estimasi parameter diperoleh melalui invers matriks dan diberikan sebagai berikut.

$$\hat{Cov}(\hat{\beta}) = \{x^T \text{Diag}[\hat{\pi}(x_i)(1 - \hat{\pi}(x_i))]x\}^{-1} \quad (2.11)$$

$$\text{dimana } X^T = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1k} & x_{2k} & \cdots & x_{nk} \end{bmatrix}.$$

$\text{Diag} [\hat{\pi}(x_i)(1 - \hat{\pi}(x_i))]$ merupakan matriks diagonal $n \times n$ dengan diagonal utama $[\hat{\pi}(x_i)(1 - \hat{\pi}(x_i))]$. Penaksir $SE(\beta)$ diberikan oleh akar kuadrat diagonal utama. Untuk mendapatkan nilai taksiran β dari turunan pertama fungsi $L(\beta)$ yang *nonlinear* maka digunakan metode iterasi *Newton Raphson*. Persamaan yang digunakan adalah sebagai berikut.

$$\beta^{(t+1)} = \beta^{(t)} - (H^{(t)})^{-1} q^{(t)}; t = 1, 2, \dots \quad (2.12)$$

dengan $q^T = \left(\frac{\partial L(\beta)}{\partial \beta_0}, \frac{\partial L(\beta)}{\partial \beta_1}, \dots, \frac{\partial L(\beta)}{\partial \beta_k} \right)$ yang merupakan vektor *Gradient* dan H merupakan matriks *Hessian*. Elemen-elemen matriks

Hessian adalah $h_{ju} = \frac{\partial^2 L(\beta)}{\partial \beta_j \partial \beta_u}$, sehingga $H = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1k} \\ h_{21} & h_{22} & \cdots & h_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ h_{k1} & h_{k2} & \cdots & h_{kk} \end{bmatrix}$.

Pada setiap iterasi berlaku hal berikut.

$$\begin{aligned} h_{ju}^{(t)} &= \frac{\partial^2 L(\beta)}{\partial \beta_j \partial \beta_u} \Big|_{\beta^{(t)}} = - \sum_{i=1}^n x_{ij} x_{iu} \pi(x_i)^{(t)} (1 - \pi(x_i)^{(t)}) \\ q_j^{(t)} &= \frac{\partial L(\beta)}{\partial \beta_j} \Big|_{\beta^{(t)}} = \sum_{i=1}^n (y_i - \pi(x_i)^{(t)}) x_{ij} \\ \pi(x_i)^{(t)} &= \frac{e^{\left(\sum_{j=0}^k \beta_j^{(t)} x_{ij} \right)}}{1 + e^{\left(\sum_{j=0}^k \beta_j^{(t)} x_{ij} \right)}} \end{aligned} \quad (2.13)$$

Dari persamaan di atas diperoleh persamaan berikut.

$$\beta^{(t+1)} = \beta^{(t)} + \{X^T \text{Diag}[\pi(x_i)^{(t)}(1 - \pi(x_i)^{(t)})]x\}^{-1} X^T(y - m^{(t)}) \quad (2.14)$$

dengan $m^{(t)} = \pi(x_i)^{(t)}$. Langkah-langkah iterasi *Newton Raphson* diberikan sebagai berikut,

1. Menentukan nilai dugaan awal $\beta^{(0)}$ kemudian dengan menggunakan persamaan 2.13 maka didapatkan $\pi(x_i)^{(0)}$.
2. Dari $\pi(x_i)^{(0)}$, pada langkah 1 diperoleh matriks *Hessian* H^0 dan vektor q^0 .
3. Proses selanjutnya untuk $t > 0$ digunakan persamaan 2.13 dan 2.14 hingga $\beta^{(t)}$ dan $\pi(x_i)^{(t)}$ konvergen.

Setelah parameter hasil estimasi diperoleh, kemudian dilakukan pengujian keberartian terhadap koefisien β secara univariat terhadap variabel respon yaitu dengan membandingkan parameter hasil maksimum *likelihood*, dugaan β dengan standard error parameter tersebut. Hipotesis pengujian parsial adalah sebagai berikut.

$$H_0: \beta_i = 0$$

$$H_1: \beta_i \neq 0; i = 1, 2, \dots, p$$

Dengan statistik uji sebagai berikut.

$$W = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \quad (2.15)$$

Statistik uji W tersebut disebut sebagai statistik uji *Wald* yang mengikuti distribusi normal sehingga H_0 ditolak jika $|W| > Z_{\alpha/2}$ dan dapat juga diperoleh melalui persamaan berikut.

$$W^2 = \frac{\hat{\beta}_i^2}{SE(\hat{\beta}_i)^2} \quad (2.16)$$

Statistik tersebut mengikuti distribusi *Chi Square* sehingga H_0 ditolak jika $W^2 > X^2_{(v, \alpha)}$ dengan v adalah *degrees of freedom* dari banyaknya prediktor dikurangi satu.

Setelah diperoleh prediktor yang signifikan berpengaruh terhadap variabel respon pengujian univariat, langkah selanjutnya adalah menentukan variabel hasil pengujian univariat mana saja yang signifikan mempengaruhi variabel respon secara bersamaan. Pengujian ini dilakukan untuk memeriksa keberartian koefisien β secara serentak terhadap variabel respon. Hipotesis yang digunakan diberikan sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_I = 0$$

$$H_1: \text{Minimal terdapat satu } \beta_i \neq 0; i = 1, 2, \dots, p$$

Dengan statistik uji sebagai berikut.

$$G = -2 \ln \frac{\left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_0}{n}\right)^{n_0}}{\sum_{i=1}^n \hat{\pi}^{y_i} (1 - \hat{\pi})^{(1-y_i)}} \quad (2.17)$$

dimana $n_1 = \sum_{i=1}^n y_i$ dan $n_0 = \sum_{i=1}^n (1 - y_i)$. Statistik uji G adalah *likelihood test* yang mengikuti distribusi *Chi Square* sehingga H_0

ditolak jika $G > X^2_{(v,\alpha)}$ dengan v adalah *degrees of freedom* dari banyaknya prediktor tanpa β_0 .

Interpretasi terhadap koefisien parameter dilakukan untuk menentukan kecenderungan fungsional antara variabel prediktor dengan variabel respon serta menunjukkan pengaruh perubahan nilai pada variabel yang bersangkutan. Dalam hal ini digunakan besaran *Odds Ratio* atau e^β yang dinyatakan dalam ψ . *Odds ratio* diartikan sebagai kecenderungan variabel respon dalam memiliki nilai tertentu jika diberikan $x = 1$ dan dibandingkan pada $x = 0$. Keputusan tidak terdapat hubungan antara variabel prediktor dengan variabel respon diambil jika nilai *odds ratio* sama dengan satu.

Jika nilai *odds ratio* kurang dari satu, maka antara variabel prediktor dan variabel respon terdapat hubungan negatif setiap kali perubahan nilai variabel bebas, dan jika *odds ratio* lebih dari satu, maka antara variabel prediktor dengan variabel respon terdapat hubungan positif setiap kali perubahan nilai variabel bebas (Agresti, 2002).

2.2 *Logistic Regression Classification by Ensembles from Random Partitions (LR CERP)*

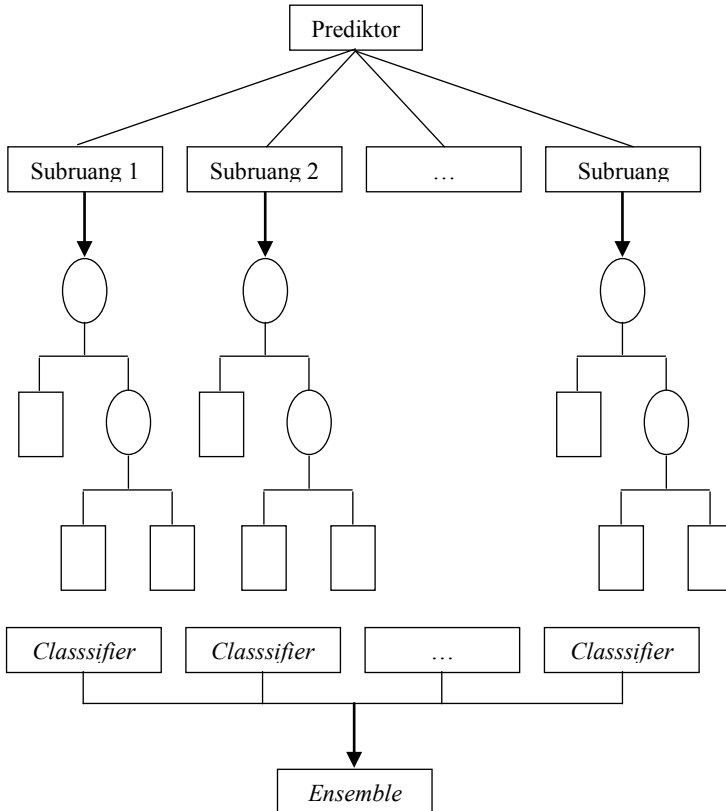
Logistic Regression CERP (LR CERP) merupakan pasangan dari *Classification Tree CERP* (C-T CERP) yang menggunakan regresi logistik sebagai *base classifier*. Misalkan Θ adalah ruang prediktor yang dipartisi secara random menjadi K subruang yang $(\theta_1, \theta_2, \dots, \theta_K)$ yang saling *mutually exclusive* dengan ukuran yang sama sehingga diasumsikan tidak terdapat bias dalam pengambilan prediktor pada tiap subruang.

Performa CERP tergantung pada banyaknya prediktor pada satu partisi yang ditentukan oleh banyaknya partisi optimal. Partisi optimal didapatkan dari persamaan berikut.

$$K = \frac{6 \times p}{n} \quad (2.18)$$

dimana p adalah banyaknya prediktor dan n adalah banyaknya pengamatan. Apabila ukuran n lebih besar dari ukuran p , maka par-

tisi optimal dapat diperoleh dengan membagi data sebanyak i menjadi $\frac{p}{i}$ dimana i adalah sembarang integer yang kurang dari n . Relatifnya partisi optimal didapatkan dari $K = \frac{p}{i}$ akurasi yang menghasilkan akurasi tertinggi. Berikut ini merupakan bagan yang menggambarkan konsep LR CERP.



Gambar 2.1 Bagan Konsep LR CERP

Suatu model klasifikasi akan dibangun pada masing-masing subruang dengan menggunakan regresi logistik yang cukup lemah dalam menangani pemilihan variabel. LR CERP meningkatkan akurasi dengan mengombinasikan masing-masing hasil

klasifikasi dari masing-masing subruang menjadi satu *ensemble*. Karena jumlah prediktor dalam satu partisi akan lebih kecil dari banyaknya pengamatan. Maka pemilihan variabel pada regresi logistik apabila jumlah prediktor lebih banyak dari jumlah pengamatan, tidak perlu dilakukan. LR CERP mengombinasikan beberapa regresi logistik untuk meningkatkan akurasi dengan mengambil rata-rata nilai prediksi dalam satu *ensemble*. Nilai prediksi dari semua *base classifiers* dirata-rata dan dikategorikan menjadi 0 atau 1 berdasarkan *threshold* (Lim, 2007).

2.3 *Logistic Regression Ensembles (LORENS)*

Berdasarkan algoritma LR CERP, Lim, Ahn, Moon, dan Chen pada tahun 2010 mengembangkan LORENS menggunakan model regresi logistik sebagai basis klasifikasi. Tujuannya adalah untuk mengombinasikan hasil model regresi logistik untuk mendapatkan satu *classifier* kuat yang dapat dibandingkan dengan metode agregasi kompleks lainnya, dalam rangka meningkatkan akurasi prediksi. LORENS mengulangi prosedur yang digunakan LR CERP beberapa kali untuk digabungkan dalam satu *ensemble* hingga terbentuk beberapa *ensemble*. Seperti LR CERP, LORENS mempartisi ruang prediktor Θ secara acak menjadi k subruang dalam ukuran yang sama, untuk meminimalkan korelasi diantara klasifikasi dalam satu ensemble. Karena subruang secara acak dipilih dari distribusi yang sama, diasumsikan bahwa tidak ada bias dalam penyeleksian variabel prediktor dalam tiap subruang. Dalam tiap subruang ini, seluruh model regresi logistik dimasukkan tanpa seleksi variabel. Melalui pengacakan ini, diharapkan probabilitas yang hampir sama dari error klasifikasi dan perbaikan akurasi yang sama pada masing-masing *classifier* dalam satu *ensemble*.

LORENS mengombinasikan nilai prediksi (*predicted value*) dari kumpulan model regresi logistik pada masing-masing partisi untuk meningkatkan akurasi dalam satu *ensemble*. Prosedur LORENS melibatkan LR CERP yang dilakukan berulang-ulang. Kombinasi tersebut dapat berupa rata-rata atau nilai terba-

nyak yang menghasilkan akurasi yang hampir sama. Namun LORENS menggunakan rata-rata karena akhirnya dapat menghasilkan akurasi yang sedikit lebih baik daripada nilai terbanyak. Nilai dari semua model dirata-rata dan diklasifikasikan menjadi 0 atau 1 dengan *threshold* tertentu. LORENS menghasilkan beberapa *ensemble* dengan partisi random yang berbeda-beda dan memilih nilai terbanyak di antara beberapa *ensemble*. Dari nilai tersebut didapatkan satu akurasi umum yang akurasinya telah ditingkatkan dari sumbangsih beberapa *ensemble* yang dibangun. Peningkatan akurasi ini relatif didapatkan bila jumlah *ensemble* yang dibangun lebih dari 10.

Threshold yang biasa digunakan dalam klasifikasi untuk respon biner adalah 0,5. Namun akurasi klasifikasi tidak akan baik apabila proporsi kelas 1 dan 0 tidak seimbang. Dalam rangka menyeimbangkan *sensitivity* dan *specificity*, LORENS mencari *threshold* optimal melalui rumus berikut.

$$Threshold = \frac{\bar{y} + 0,5}{2} \quad (2.19)$$

dengan \bar{y} adalah probabilitas pengamatan berada di kelas positif. Tahapan dalam pengklasifikasian dilakukan sebagai berikut.

1. Membentuk model logit dari data *training*.
2. Mensubstitusikan data *testing* ke dalam model logit hingga mendapatkan nilai probabilitas.
3. Mengklasifikasikan pengamatan data *testing* ke dalam kelas positif jika nilai probabilitas pengamatan lebih besar dari *threshold*, dan mengklasifikasikan pengamatan ke dalam kelas negatif jika nilai probabilitas yang didapatkan lebih kecil dari nilai *threshold*.
4. Membandingkan prediksi klasifikasi dengan kelas aktual.
5. Mengelompokkan hasil perbandingan ke dalam kelompok *TP*, *TN*, *FP*, dan *FN*.

dengan *TP* (*True Positive*) adalah total konsumen kelas positif yang terprediksi tepat di kelas positif dan *TN* (*True Negative*) adalah total konsumen negatif yang terprediksi tepat di kelas ne-

gatif. Sedangkan *FP* (*False Positive*) adalah total konsumen negatif yang salah terprediksi menjadi positif dan *FN* (*False Negative*) adalah total konsumen positif yang salah terprediksi menjadi negatif.

Jika jumlah kelas adalah dua, tabel berikut menunjukkan prediksi klasifikasi dan klasifikasi aktual.

Tabel 2.1 Tabulasi Silang Klasifikasi Prediksi dan Klasifikasi Sebenarnya

		Kelas Aktual	
		$p (+)$	$n (-)$
Kelas Prediksi	$p (+)$	<i>True Positive</i> (TP)	<i>False Positive</i> (FP)
	$n (-)$	<i>False Negative</i> (FN)	<i>True Negative</i> (TN)

Ketepatan prediksi klasifikasi dihitung dari jumlah prediksi yang tepat dibagi dengan total jumlah prediksi. Catal (2012) menghitung ukuran ketepatan klasifikasi tersebut sebagai berikut.

$$\text{Sensitivity} = \frac{TP}{(TP+FN)} \quad (2.20)$$

$$\text{Specificity} = \frac{TN}{(FP+TN)} \quad (2.21)$$

$$\text{Akurasi} = \frac{(TP+TN)}{(TP+FP+TN+FN)} \quad (2.22)$$

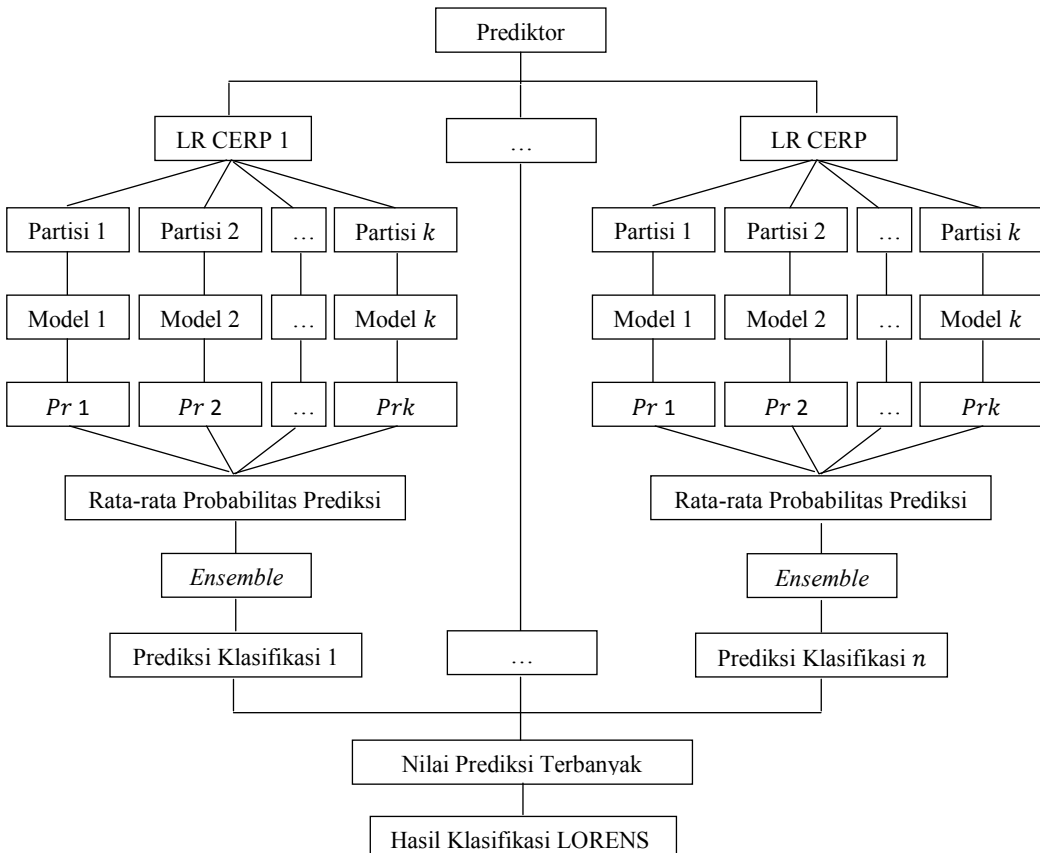
Sementara itu, kesalahan klasifikasi dapat dinyatakan dalam *False Positive Rate* (FPR), *False Negative Rate* (FNR), dan *Error*. Berikut ini merupakan formula untuk mendapatkan ukuran kesalahan klasifikasi tersebut.

$$\text{FPR} = \frac{FP}{(FP+TN)} \quad (2.23)$$

$$\text{FNR} = \frac{FN}{(TP+FN)} \quad (2.24)$$

$$\text{Error} = \frac{(FP+FN)}{(TP+FP+TN+FN)} \quad (2.25)$$

Sehingga FPR adalah $1 - \text{Specificity}$, FNR adalah $1 - \text{Sensitivity}$, dan *Error* adalah $1 - \text{Akurasi}$. Gambar di bawah ini merupakan diagram yang menggambarkan konsep LORENS.



Gambar 2.2 Diagram Alir LORENS

Karena partisi acak, LORENS bebas dari masalah dimensi data. Pemilihan variabel tidak diperlukan di LORENS dan korelasi antar klasifikasi dapat dikurangi melalui partisi acak. Dari pada LR CERP, LORENS unggul dalam basis klasifikasi yang menggunakan regresi logistik yang populer dan mudah dipahami. Selain itu LORENS juga lebih efisien dalam hal komputasi dibandingkan LR CERP yang menggunakan algoritma pohon (*tree*). Dengan mengintegrasikan keuntungan ini menuju pemili-

han *ensemble* umum, keakuratan metode dapat menjadi lebih baik (Lee dkk, 2013).

2.4 *Holdout dan Cross Validation*

Witten, Frank, dan Hall (2011) telah menjabarkan beberapa prosedur untuk mengevaluasi performa model dalam memprediksi melalui data *training* dan *testing*, diantaranya adalah *Holdout* dan *cross Validation*. Metode *holdout* mengambil sejumlah data untuk *testing* dan menggunakan sisanya untuk *testing*. Dalam prakteknya, sepertiga bagian dari data digunakan untuk *testing* dan dua-pertiga sisanya digunakan untuk *training*. Sampel yang digunakan untuk *training* (atau *testing*) mungkin tidak representatif.

Ada satu cek sederhana yang memeriksa apakah sampel representatif atau tidak yaitu setiap kelas dalam dataset penuh harus terwakili dalam proporsi yang tepat untuk *training* dan *testing*. Jika semua sampel dengan kelas tertentu dihilangkan dari *training set*, *classifier* tidak dapat diharapkan belajar dengan baik dari data yang tersedia dalam melakukan klasifikasi pada *testing set*. Sebaliknya, harus dipastikan bahwa pengambilan sampel dilakukan dengan cara random yang menjamin bahwa setiap kelas diwakili benar baik pada *training* dan *testing set*. Berikut ini merupakan langkah sederhana dalam stratifikasi.

1. Memisahkan data berdasarkan kelas tertentu.
2. Mengambil sampel dari masing-masing kelas dengan proporsi yang sama yang dikehendaki.
3. Menggabungkan sampel dari dua kelas yang sudah diambil.

Prosedur ini disebut stratifikasi yang apabila dikombinasikan dengan prosedur *holdout*, maka akan menciptakan *stratified holdout*. Meskipun umumnya layak dilakukan, stratifikasi hanya menyediakan perlindungan yang lemah terhadap kelas yang tidak representatif dalam *training* dan *testingset*. Cara yang lebih umum untuk mengurangi bias yang terjadi pada pengambilan sampel dengan *holdout* adalah dengan mengulang seluruh proses pada

data *training* dan *testing* beberapa kali menggunakan sampel acak yang berbeda. Dalam setiap iterasi proporsi tertentu, sebagian besar dari data dipilih secara acak untuk *training* yang sebelumnya telah dilakukan stratifikasi dan sisanya digunakan untuk *testing*. Tingkat kesalahan pada iterasi yang berbeda dirata-ratakan untuk menghasilkan tingkat kesalahan keseluruhan. Metode ini secara sederhana dikenal sebagai metode *holdout* berulang untuk mengestimasi tingkat kesalahan.

Berikut ini adalah langkah-langkah yang dilakukan dalam metode *Holdout*.

1. Memisahkan data berdasarkan kelas respon
2. Mengambil secara acak sebagian pengamatan pada masing-masing kelas sebanyak *testing set* yang dibutuhkan.
3. Menggabungkan masing-masing *testing set* yang telah terambil dari masing-masing kelas sebagai data *testing*.
4. Menggabungkan sisa pengamatan dari dua kelas menjadi data *training*

Dalam prosedur *holdout* tunggal, penukaran peran data *training* dan *testing* dapat dipertimbangkan untuk melatih sistem dan hasil dari dua proses tersebut dirata-rata untuk mengurangi efek dari data yang tidak representatif pada data *training* dan *testing*. Sayangnya, hal ini hanya dapat dilakukan dengan membagi data menjadi 50:50 antara *training* dan *testing* set, yang idealnya lebih baik menggunakan lebih dari setengah data untuk *training*. Maka dari itu terbentuklah *cross validation*.

Dalam *cross validation* data dibagi menjadi beberapa k *folds* atau partisi yang sama banyak, masing-masing pada gilirannya digunakan untuk *testing* dan sisanya digunakan untuk *training*. Artinya, menggunakan satu partisi dari data untuk *testing* dan $k - 1$ *folds* sisanya digunakan untuk *training*. Prosedur ini diulangi hingga semua partisi telah diperlakukan sebagai *testing* set. Prosedur Ini disebut sebagai k *folds cross validation*, dan apabila stratifikasi juga diadopsi maka prosedur ini disebut stratified k *folds cross validation*.

Cara standar memprediksi tingkat kesalahan adalah dengan menggunakan stratifikasi 10 *folds cross validation*. Data dibagi secara acak menjadi 10 bagian di mana kelas diwakili dengan proporsi yang sama seperti dalam dataset yang penuh. Dengan demikian, prosedur dijalankan 10 kali di set *training* yang berbeda (setiap set memiliki banyak anggota yang sama dengan set lainnya). Akhirnya, 10 perkiraan kesalahan dirata-ratakan untuk menghasilkan perkiraan kesalahan secara keseluruhan. Tes pada berbagai dataset yang berbeda dengan teknik belajar yang berbeda telah menunjukkan bahwa 10 adalah yang tepat dari lipatan untuk mendapatkan estimasi kesalahan yang terbaik. Beberapa bukti teori yang mendukung hal ini. Meskipun argumen ini tidak berarti konklusif di kalangan *Machine Learning* dan *Data Mining* tentang apa yang skema terbaik untuk evaluasi, 10 *foldscross validation* telah menjadi metode standar dalam praktis. Tes juga menunjukkan bahwa penggunaan stratifikasi meningkatkan ketepatan prediksi. Berikut ini merupakan langkah-langkah yang dilakukan dalam metode *Cross Validation*.

1. Memisahkan pengamatan berdasarkan kelas respon.
2. Membagi pengamatan menjadi 10 bagian di masing-masing kelas.
3. Menggabungkan bagian yang sama dari kedua kelas.
4. Menggunakan bagian 1 sebagai data *testing* dan bagian 2 sampai 10 sebagai data *training* pada *fold* pertama. Penggunaan bagian ke-2 hingga ke-10 sebagai data *testing* dilakukan di *fold* 2 hingga *fold* 10 secara berurutan.

Sepuluh *folds cross validation* mungkin tidak cukup mampu mendapatkan estimasi kesalahan yang handal dan konsisten. Stratifikasi mengurangi variasi, tapi tentu tidak menghilangkannya sama sekali. Ketika mencari perkiraan kesalahan yang akurat, adalah prosedur standar untuk mengulangi proses *cross validation* 10 kali dan merata-rata hasilnya. Prosedur ini melibatkan 100 kali algoritma pada dataset, dalam rangka mendapatkan performa ukuran yang baik memang membutuhkan usaha komputasional

yang intensif. Pembagian data menjadi $\frac{9}{10}$ bagian sebagai data *training* dan $\frac{1}{10}$ bagian sebagai data *testing*, dapat diadopsi pada prosedur *holdout* (Witten dkk, 2011).

2.5 Pembelotan Konsumen Perusahaan 'X'

Cloud Based Software berbeda dari aplikasi atau perangkat lunak biasa. Secara definisi, *Cloud* adalah sistem komputasi suatu layanan di internet yang dapat dilakukan melalui *personal computer* dan didukung oleh jaringan internet atau jaringan lokal (*local network*). Layanan tersebut bisa berupa penyimpanan, jejaring sosial, layanan *e-mail* dan perangkat lunak lainnya. Mayoritas penelitian-penelitian pendeteksian pembelotan pelanggan sebelumnya dilakukan pada perusahaan telekomunikasi yang masing-masing perangkat konsumen hanya dapat terhubung dengan satu *provider*. Sedangkan *Cloud Based Software* memungkinkan satu konsumen memiliki lebih dari satu akun layanan *web based hosting* dan terhubung dengan beberapa *provider* sekaligus. Dalam kasus yang dihadapi Perusahaan „X“, seorang konsumen dapat memiliki lebih dari satu program antivirus yang dapat aktif digunakan secara bersamaan dalam satu perangkat.

Berdasarkan harga, Perusahaan „X“ membagi produknya menjadi tiga kelas yaitu produk *Low Price*, *Medium Price*, dan *High Price*. Perusahaan „X“ memiliki situs *e-commerce* yang mengirim *e-mail* konfirmasi *auto renewal* (pembaharuan/ perpanjangan kontrak secara otomatis) suatu produk kepada konsumen minimal dua kali dalam satu tahun yaitu pada hari pertama setelah kontrak disetujui dan 50 hari sebelum masa kontrak selesai.

Di dalam *e-mail* tersebut, konsumen dihadapkan pada pilihan „*opt-in*“ dan „*opt-out*“. Jika konsumen memilih „*opt-in*“, maka konsumen akan mengisi form pembaharuan kontrak. Jika konsumen memilih „*opt-out*“, maka konsumen membelot atau tidak lagi menggunakan produk tersebut. Dalam pelaksanaannya, konsumen dapat memilih „*opt-out*“ untuk mengingggalkan suatu produk dan mulai berlangganan produk lain yang ditawarkan Perusahaan „X“.

Dalam kondisi ini, konsumen tersebut tidak benar-benar membelot. Sementara itu, situs *e-commerce* hanya mencatat aktifitas „*opt-out*“ dan „*opt-in*“ konsumen tanpa memeriksa aktifitas lanjutan konsumen tersebut di Perusahaan „X“.

Maka dari itu, *data preparation* sangat penting dilakukan. Catatan asli yang direkam situs *e-commerce* Perusahaan „X“ memuat 30 variabel. Martono (2014) telah melakukan *preprocessing* terhadap variabel tersebut melalui beberapa perlakuan seperti berikut.

1. Jika terdapat variabel-variabel yang memiliki nilai yang sama dan/atau definisi yang berkaitan, maka variabel yang tepat dipilih untuk digunakan.
2. Variabel yang tidak memiliki relasi dengan proses membeli dan membelot tidak digunakan.
3. Variabel yang memberikan nilai sama untuk semua konsumen tidak digunakan.
4. Variabel kualitatif yang tidak dapat ditransformasi tidak digunakan.

Tabel di bawah ini menampilkan sepuluh variabel terpilih yang masih membutuhkan transformasi lebih lanjut. Penamaan variabel yang ditabelkan berikut merupakan istilah asli dari penelitian tersebut berdasarkan catatan di situs *e-commerce*.

Tabel 2.2 Variabel Asli dari Situs *E-commerce* Perusahaan „X“

Variabel	Definisi
AR_KEY	<i>Serial Key</i> produk
AR_FLAG	Konfirmasi pembaharuan terakhir
OPTIN_DATE	Tanggal pembaharuan kontrak
AR_COUNT	Total pembaharuan
CC_PRODUCT_BASE	<i>Base</i> produk
CC_PRODUCT_PRICE	Harga produk
CC_SUBS_DAY	Periode valid produk
OPTIONAL_FLAG	Konfirmasi penggunaan layanan pilihan
ORG_FLAG	Tipe konsumen
MAIL_STATUS	Status pengiriman <i>e-mail</i> pembaharuan

Beberapa fitur baru harus diekstrak dari data asli *purchase* and *autorenewal* untuk memprediksi pembelotan yang sebenar-

nya. Data di atas memuat pola pembatalan pembelotan yaitu ketika konsumen memilih „*opt-out*“ dari suatu kontrak lalu melakukan „*opt-in*“ ke pilihan produk lain. Prasasti dkk (2013) melakukan transformasi variabel pada data tersebut, namun penjelasan transformasi tersebut dijelaskan oleh Martono dkk (2014). Transformasi membentuk variabel baru yang benar-benar merepresentasikan konsumen yang membelot dan tidak membelot melalui aturan-aturan berikut.

1. Jika konsumen memilih „*opt-out*“, maka AR_FLAG=1. Namun jika konsumen dengan AR_KEY dan CC_PRODUCT_BASE yang sama tercatat „*opt-in*“, maka konsumen tidak dikategorikan membelot.
2. Jika AR_FLAG=0, maka konsumen membelot.
3. Selain poin 1 dan 2, konsumen tidak membelot.

Transformasi yang dilakukan Niken menghasilkan dua variabel baru yaitu UPDATE COUNT dan CLASS. UPDATE COUNT merupakan total pembelian dan pembaharuan selain pembelian pertama yang dilakukan konsumen, sedangkan variabel CLASS mendefinisikan konsumen sebagai membelot dan tidak membelot. Dari tahap *preprocessing* ini, Niken mendapatkan lima variabel yang digunakan dalam penelitian ini. Variabel-variabel tersebut adalah UPDATE_COUNT (Akumulasi Update), CC_PRODUCT_PRICE (Harga Produk), VSSA_FLAG (Jawaban Kontrak), ORG_FLAG (Tipe Konsumen), MAIL_STATUS (Status Pengiriman), dan RIHAN_FLAG (Pembelotan Konsumen).

2.6 Penelitian Sebelumnya

Martono dkk (2014) sebelumnya menggunakan dua jenis data yaitu data *purchase and autorenewal* dan data dari aktifitas web log (*web log*). Kedua data tersebut dibandingkan untuk prediksi pembelotan pelanggan yang lebih baik. Data *purchase and autorenewal* berisi catatan aktivitas pelanggan dalam membeli dan memperbaharui produk mereka selama enam tahun (dari tahun 2007 sampai dengan 2013).

Data *purchase and autorenewal* mencakup identitas kontrak pelanggan, status terbaru dari bendera pembaharuan, tanggal terbaru *autorenewal* kontrak, jumlah pembelian dan perpanjangan, jenis produk dasar bahwa pelanggan dibeli, total pembayaran oleh pelanggan, periode surat perintah produk, apakah layanan opsional yang digunakan, jenis pelanggan (personal atau komersial), dan status pengiriman *e-mail*.

Sama seperti data *purchase and autorenewal*, data web *log* juga diambil dari situs *e-commerce* perusahaan „X“. Data ini mencakup file *log* enam bulan (Januari-Juni 2013) yang mengandung jumlah pembayaran oleh pelanggan, penggunaan layanan opsional, jenis pelanggan (personal atau komersial), jenis sistem operasi menggunakan pelanggan, jenis *browser* pelanggan digunakan untuk *browsing* internet, jumlah tampilan halaman *website*, jumlah situs kunjungan, jumlah tampilan produk dan pandangan keranjang, dan jumlah pesanan pelanggan telah dibuat.

Dalam tiga penelitian sebelumnya Prasasti dkk (2013), Prasasti dan Ohwada (2014), dan Martono dkk (2014) menggunakan teknik *Machine Learning* dengan beberapa metode klasifikasi seperti *Decision Tree*, SVM, *Random Forest* dan MPO. Berikut ini adalah akurasi dari hasil analisis penelitian-penelitian tersebut.

Tabel 2.3 Perbedaan Akurasi dari Penelitian Sebelumnya

Machine Learning	Low Price	Medium Price	High Price	Peneliti
C4.5	0,7212	0,8195	0,8287	Prasasti dkk (2013)
SVM	0,8261	0,7268	0,8294	Prasasti dkk (2013)
J48	0,7212	0,8195	0,8287	Prasasti dan Ohwada (2014)
RF	0,7228	0,8232	0,8313	Prasasti dan Ohwada (2014)
MLP	0,6881	0,7873	0,6857	Prasasti dan Ohwada (2014)
SMO	0,6881	0,8228	0,8271	Prasasti dan Ohwada (2014)
C4.5	0,828	0,724	0,833	Martono dkk (2014)

Prasasti dkk (2013) memprediksi pembelotan konsumen Perusahaan „X” menggunakan beberapa metode klasifikasi yaitu *C4.5 Decision Tree* dan *Support Vector Machine* (SVM) yang keduanya memberikan akurasi tertinggi pada produk *High Price*. Pada tahun 2014 Prasasti dan Ohwada meneliti kembali kasus ini dengan metode klasifikasi *J48 Decision Tree* (J48), *Random Forest* (RF), *Neural Network* dengan fungsi *Multi Layer Perception* (MLP), dan *Support Vector Machine* algoritma SMO. J48, RF dan SMO menghasilkan akurasi terbaik pada produk *High Price*, sedangkan MLP memberikan hasil terbaik pada produk *Medium Price*. Pada tahun yang sama Martono dkk (2014) melanjutkan penelitian ini dengan metode klasifikasi *C4.5 Decision Tree* yang menghasilkan akurasi terbaik pada produk *High Price*.

Ketiga penelitian di atas dilakukan pada kasus yang sama namun dengan data yang berbeda. Performa *classifiers* terlihat berbeda-beda terhadap masing-masing jenis produk. Akurasi tertinggi produk *Low Price* dan *High Price* dihasilkan oleh C4.5, sedangkan untuk produk *Medium Price* dihasilkan oleh *Random Forest*.

Halaman ini sengaja dikosongkan.

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder dari penelitian yang dilakukan oleh Prasasti dkk (2013), Prasasti dan Ohwada (2014), dan Martono dkk (2014), yang telah melalui *preprocessing data*. Sedangkan data asli disediakan oleh perusahaan penyedia software antivirus berbasis internet Perusahaan 'X' yang diambil dari situs *e-commerce* perusahaan tersebut dari tahun 2007 hingga tahun 2013. Data penelitian ini adalah catatan dari aktifitas konsumen terhadap produk Perusahaan 'X' yang dibedakan berdasarkan harganya, yaitu produk *Low Price*, *Medium Price*, dan *High Price*. Unit pengamatan dalam data tersebut adalah konsumen Perusahaan 'X' dengan ukuran sampel sebanyak 500000 konsumen untuk produk *Low Price*, 408810 konsumen untuk produk *Medium Price*, dan 709899 konsumen untuk produk *High Price*.

3.2 Variabel Penelitian

Variabel penelitian yang digunakan dalam penelitian ini terdiri dari satu variabel respon (Y) dan lima variabel prediktor (X) dari transformasi yang dilakukan oleh Prasasti dkk (2013). Berikut ini merupakan penjabaran dari variabel-variabel tersebut.

1. Variabel AKUMULASI UPDATE (X_1)

Variabel Akumulasi Update merupakan akumulasi banyaknya *update* dari pembelian dan pembaharuan. Setiap kali konsumen melakukan pembelian atau pembaharuan maka Akumulasi Update akan tercatat bertambah 1. Pengamatan Akumulasi Update dalam data tercatat adalah mulai dari 0 sampai tidak terbatas..

2. Variabel HARGA PRODUK (X_2)

Variabel Harga Produk merupakan harga produk yang baru saja dibeli yang berkisar dari 1886 hingga 39000 Yen Jepang (JPY).

3. Variabel JAWABAN KONTRAK (X_3)

Variabel Jawaban Kontrak merupakan pilihan konsumen untuk melanjutkan atau menghentikan kontrak dengan nilai 1 untuk ‘*opt-in*’ (melanjutkan menggunakan produk tertentu) dan 0 untuk ‘*opt-out*’ (berhenti menggunakan produk tertentu).

4. Variabel TIPE KONSUMEN (X_4)

Variabel Tipe Konsumen merupakan tipe konsumen dengan nilai 0 untuk individu dan 1 untuk organisasi.

5. Variabel STATUS PENGIRIMAN (X_5)

Perusahaan ‘X’ mengirim *e-mail* tawaran perpanjangan kontrak. Variabel Status Pengiriman merupakan status pengiriman *e-mail* yang bernilai 1 bila terkirim dan 0 bila tidak terkirim.

6. Variabel PEMBELOTAN KONSUMEN (Y)

Variabel Pembelotan Konsumen merupakan klasifikasi konsumen yang bernilai 1 jika membelot dan bernilai 0 jika konsumen lanjut menggunakan produk satu atau lebih produk antivirus Perusahaan ‘X’ meskipun untuk produk yang berbeda.

Berikut ini tabel yang memuat struktur data, variabel penelitian berdasarkan fitur *e-commerce*, skala variabel, jenis bilangan serta besaran nilai yang diamati.

Tabel 3.1 Struktur Data Penelitian

Variabel	X_1	X_2	X_3	X_4	X_5	Y
Fitur	Akumulasi Update	Harga Produk	Jawaban Kontrak	Tipe Konsumen	Status Pengiriman	Pembelotan Konsumen
	x_{11}	x_{21}	x_{31}	x_{41}	x_{51}	y_1
Struktur Data	x_{12}	x_{22}	x_{32}	x_{42}	x_{52}	y_2
	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
	x_{1n}	x_{2n}	x_{3n}	x_{4n}	x_{5n}	y_n
Jenis	Kontinyu	Kontinyu	Kategorik	Kategorik	Kategorik	Kategorik

3.3 Langkah Analisis

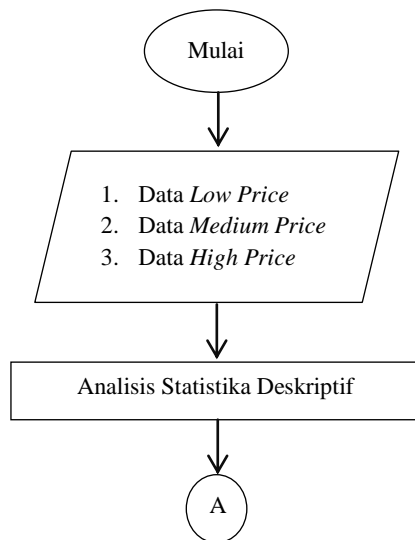
Dalam penelitian ini akan dilakukan analisis klasifikasi dengan *Binary Logistic Regression* dan LORENS yang keduanya dievaluasi dengan metode *Holdout*, serta analisis klasifikasi dengan LORENS yang dievaluasi dengan metode *Cross Validation*. Langkah analisis yang sama dilakukan pada data *Low Price*, *Medium Price*, dan *High Price*. Langkah-langkah analisis untuk semua jenis data adalah sebagai berikut:

1. Membuat analisa deskriptif terhadap tiap variabel.
 - a. Menghitung *mean* dan *varians* untuk data bertipe kontinyu.
 - b. Menghitung *sum* tiap kategori untuk data bertipe kategorik.
2. Melakukan analisis klasifikasi dengan *Binary Logistic Regression* dan prosedur evaluasi *Holdout*.
 - a. Melakukan stratifikasi.
 - b. Mengambil sampel 10% secara random dari total data terstratifikasi sebagai data testing serta menggunakan 90% data lainnya sebagai data *training*.
 - c. Mengestimasi parameter dengan *maximum likelihood* dan iterasi *Newton Raphson* dari data *training*.
 - d. Menyusun model regresi logistik terbaik.
 - e. Melakukan transformasi logit model regresi logistik.
 - f. Membentuk persamaan linear.
 - g. Menguji signifikansi parameter secara parsial.
 - h. Menginterpretasi parameter dan *Odd Ratio*.
 - i. Memprediksi klasifikasi dari data testing.
 - j. Menghitung nilai akurasi, *sensitivity*, dan *specificity*.
3. Melakukan analisis klasifikasi dengan LORENS dengan prosedur evaluasi *Holdout*.
 - a. Melakukan stratifikasi.
 - b. Mengambil sampel 10% secara random dari total data terstratifikasi sebagai data testing serta menggunakan 90% data lainnya sebagai data *training*.

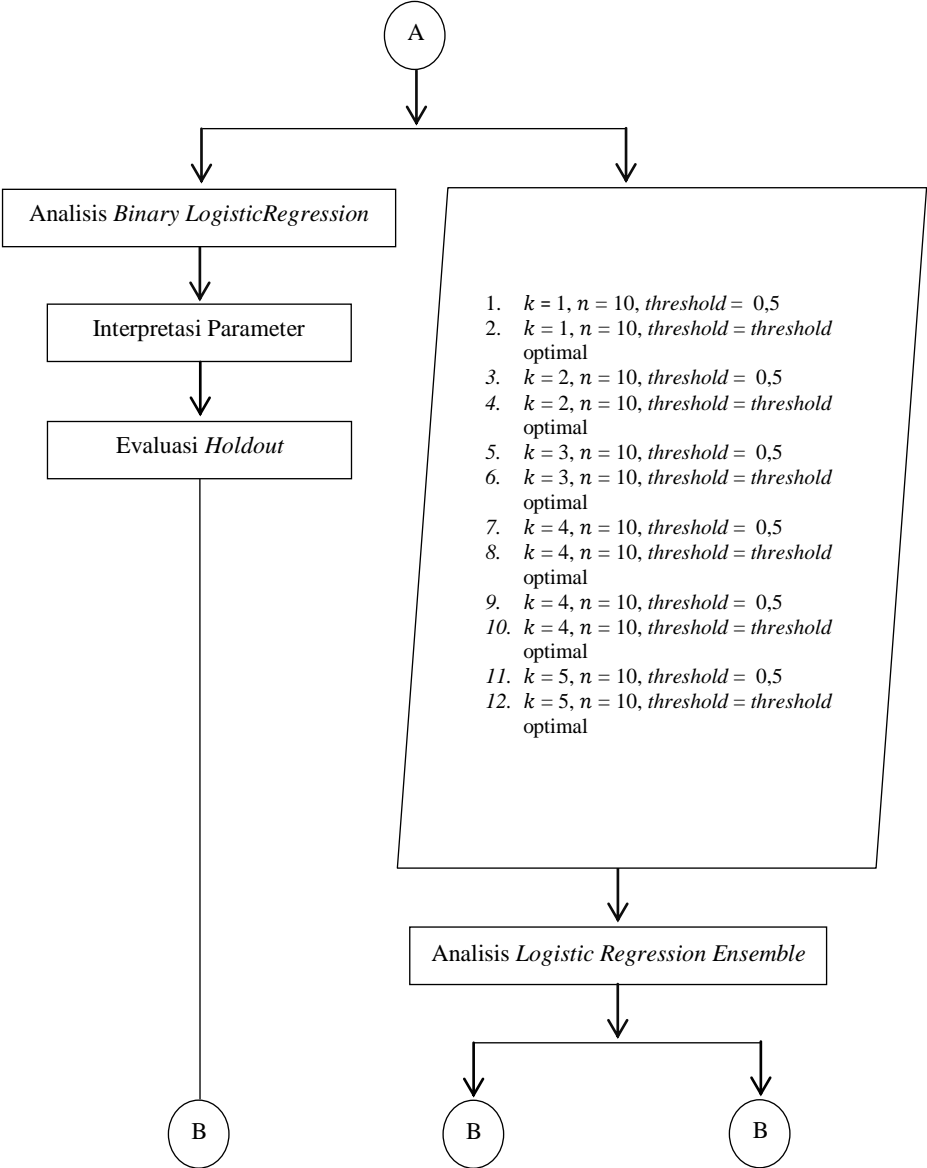
- c. Menentukan banyak partisi (k) serta banyak *ensemble* (n) dimana $k = 1, 2, \dots, 5$ dan $n = 10$.
 - d. Mempartisi variabel prediktor menjadi k subruang untuk satu *ensemble*.
 - e. Menyusun model LR masing-masing subruang partisi dari data *training*.
 - f. Mendapatkan nilai prediksi dari masing-masing model untuk semua pengamatan dari data *testing*.
 - g. Menghitung rata-rata dari semua nilai prediksi untuk masing-masing pengamatan.
 - h. Mengulangi langkah (a) hingga (e) hingga terbentuk *nensemble*.
 - i. Mencari nilai prediksi terbanyak masing-masing pengamatan di antara semua *ensemble*.
 - j. Menghitung nilai *threshold* optimal.
 - k. Membandingkan hasil dari langkah (g) dengan nilai *threshold* 0,5 dan *threshold* optimal untuk mengklasifikasikan pengamatan.
 - l. Menghitung nilai akurasi, *sensitivity*, dan *specificity*.
4. Melakukan analisis klasifikasi dengan LORENS dengan prosedur evaluasi *Cross Validation*.
- a. Melakukan stratifikasi
 - b. Membagi data menjadi 10 bagian.
 - c. Mengambil sampel 1 bagian data sebagai data *testing* dan menggunakan 9 bagian data lainnya sebagai data *training*.
 - d. Menentukan banyak partisi (k) serta banyak *ensemble* (n) dimana $k = 1, 2, \dots, 5$ dan $n = 10$.
 - e. Mempartisi variabel prediktor menjadi k subruang untuk satu *ensemble*.
 - f. Menyusun model LR masing-masing subruang partisi dari data *training*.
 - g. Mendapatkan nilai prediksi dari masing-masing model untuk semua pengamatan dari data *testing*.

- h. Menghitung rata-rata dari semua nilai prediksi untuk masing-masing pengamatan.
 - i. Mengulangi langkah (a) hingga (e) hingga terbentuk *n ensemble*.
 - j. Mencari nilai prediksi terbanyak masing-masing pengamatan di antara semua *ensemble*.
 - k. Menghitung nilai *threshold* optimal.
 - l. Membandingkan hasil dari langkah (g) dengan nilai *threshold* 0,5 dan *threshold* optimal untuk mengklasifikasikan pengamatan.
 - m. Mengulangi langkah (a) hingga (j) hingga semua bagian telah diperlakukan dalam *training* dan *testing*.
 - n. Menghitung nilai akurasi, *sensitivity*, dan *specificity*.
5. Memilih metode terbaik dari hasil analisis pada langkah 2 dan 3 yang menghasilkan ketepatan klasifikasi terbaik.
 6. Membuat kesimpulan dari hasil analisis.

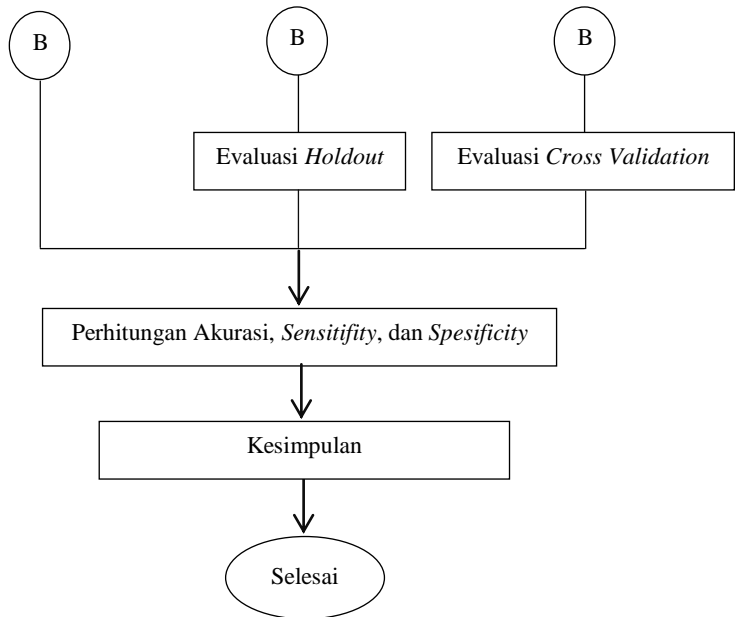
Diagram alir (*flow chart*) analisis dalam penelitian ini secara umum diilustrasikan seperti pada gambar berikut.



Gambar 3.1 Diagram Alir Penelitian



Gambar 3.2 Diagram Alir Penelitian (Lanjutan Pertama)



Gambar 3.3 Diagram Alir Penelitian (Lanjutan Kedua)

Halaman ini sengaja dikosongkan.

BAB IV ANALISIS DAN PEMBAHASAN

Analisis dan pembahasan pada penelitian ini dilakukan terhadap data catatan dari aktifitas konsumen Perusahaan ‘X’ yang terdiri atas variabel prediktor Akumulasi Update (X_1), Harga Produk (X_2), Jawaban Kontrak (X_3), Tipe Konsumen (X_4), Status Pengiriman (X_5), dan variabel respon Pembelotan Konsumen (Y). Analisis dan pembahasan dilakukan terhadap tiga jenis konsumen yang dibedakan berdasarkan harga produknya yaitu *Low Price*, *Medium Price*, dan *High Price*. Analisis dan pembahasan dalam penelitian ini akan meliputi statistika deskriptif, *Binary Logistic Regression* dan LORENS. Selain itu LORENS dengan *Cross Validation* akan ditambahkan.

4.1 Analisis Statistika Deskriptif

Analisis statistika deskriptif dalam penelitian ini dilakukan terhadap semua variabel dalam pengamatan. Analisis ini meliputi *mean* dan *standard deviation* untuk variabel kontinyu serta *sum* untuk variabel kategorik pada masing-masing kategori dalam diagram batang. Selain itu statistik deskriptif lain seperti *median* dan modus.

Terdapat 500000 konsumen yang diamati dalam data *Low Price*, 408810 dalam data *Medium Price*, dan 709899 konsumen pada data *High Price*. Tabel berikut ini menampilkan statistik deskriptif untuk variabel Harga Produk.

Tabel 4.1.1 Statistik Deskriptif Variabel Harga Produk

Harga Produk	Mean	St. Dev.	Min	Median	Max	Mode
<i>Low Price</i>	6593,50	2340,40	1886	5700	39000	5700
<i>Medium Price</i>	5323,30	1865,50	2350	4500	24000	4500
<i>High Price</i>	8932,60	3261,50	1886	7600	29981	7500

Dari tabel di atas, diketahui bahwa rata-rata pengeluaran terakhir konsumen produk antivirus *Low Price* adalah sebesar 6590,5 JPY (Jepang Yen) dengan deviasi standar sebesar 2340,4 JPY dalam setahun. Rata-rata pengeluaran terakhir konsumen

Medium Price adalah sebesar 5323,3 JPY dengan deviasi standar sebesar 1865,5 JPY. Sedangkan rata-rata pengeluaran terakhir konsumen *High Price* adalah sebesar 8932,6 JPY dengan deviasi standar sebesar 3261,5 JPY.

Pengeluaran terkecil yang pernah dilakukan terakhir kali konsumen *Low Price* adalah sebesar 1886 JPY, sedangkan pengeluaran terbesar terakhir kali yang pernah dilakukan adalah sebesar 39000 JPY. Untuk konsumen *Medium Price*, pengeluaran terkecil yang pernah dilakukan terakhir kali adalah sebesar 2350 JPY, sedangkan pengeluaran terbanyak yang pernah dilakukan mereka adalah 24000 JPY. Sedangkan pengeluaran terkecil konsumen *High Price* adalah sebesar 1886 JPY dan pengeluaran terbanyak mereka adalah sebesar 29981 JPY

Mayoritas konsumen terakhir kali mengeluarkan 5700 JPY untuk membeli produk antivirus *Low Price* dan mayoritas konsumen *Medium Price* melakukan pengeluaran terakhir sebesar 4500 JPY. Sedangkan sebagian besar konsumen *High Price* mengeluarkan 7500 JPY untuk produk antivirus *High Price*. Variabel lain selain Harga Produk bukanlah variabel kontinyu, sehingga variabel-variabel lainnya ditampilkan dalam bentuk diagram batang agar terlihat nilai minimum, maksimum, dan nilai-nilai yang sering muncul.

Berikut ini merupakan statistika deskriptif yang menggambarkan frekuensi *update* produk yang dilakukan konsumen-konsumen produk antivirus *Low Price*. *Update* tersebut dapat berupa perpanjangan dan pembaharuan produk.

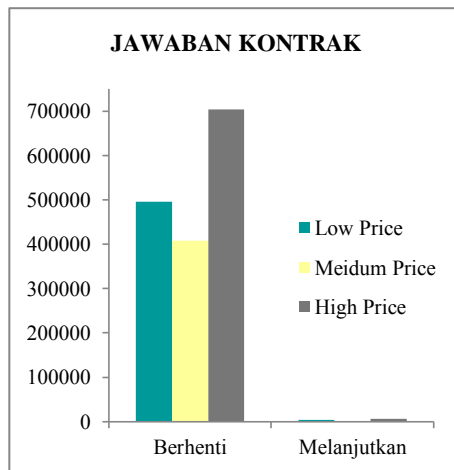
Tabel 4.1.2 Statistik Deskriptif Variabel Akumulasi Produk

Harga Produk	Mean	St. Dev.	Min	Median	Max	Mode
<i>Low Price</i>	1,399656	1,5244183	0	1	7	0
<i>Medium Price</i>	1,8112962	1,6380429	0	1	7	0
<i>High Price</i>	1,4529701	1,52454072	0	1	6	0

Berdasarkan tabel di atas, diketahui bahwa mayoritas konsumen produk *Low Price*, *Medium Price*, dan *High Price* tidak melakukan pembaharuan produk apapun, hal tersebut ditunjukkan oleh statistik modus yang bernilai 0. Pada konsumen *Low*

Price dan *Medium Price*, paling sedikit di antara mereka tidak melakukan *update* sama sekali, namun ada pula yang paling banyak melakukan *update* hingga tujuh kali. Pada konsumen *High Price*, *update* terbanyak yang pernah dilakukan konsumen adalah hingga tujuh kali. Rata-rata konsumen *Low Price* dan *High Price* melakukan *update* 1 kali sedangkan konsumen *Medium Price* rata-rata melakukan *update* 2 kali.

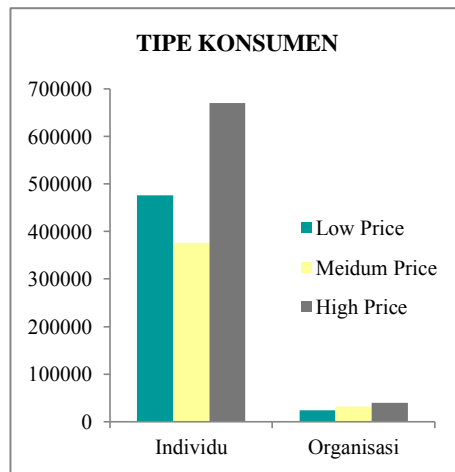
Berikut ini merupakan diagram batang untuk variabel Jawaban Kontrak. Variabel ini merujuk jawaban konsumen pada tawaran perpanjangan kontrak penggunaan produk yang dikirim perusahaan 'X' melalui *e-mail*.



Gambar 4.1 Diagram Batang Jawaban Kontrak

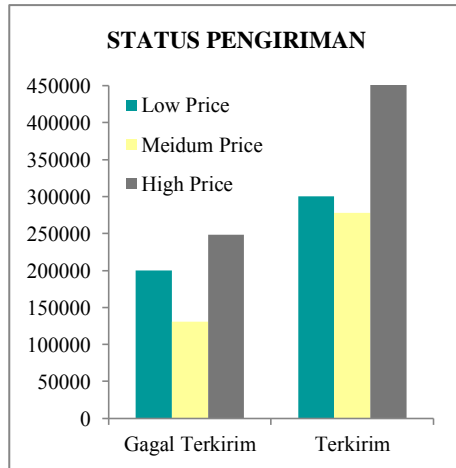
Di antara semua konsumen, sebagian besar menjawab tawaran perpanjangan kontrak dengan menjawab berhenti menggunakan produk antivirus perusahaan 'X'. Hanya terdapat 4013 konsumen *Low Price*, 996 konsumen *Medium Price*, dan 6067 konsumen *High Price* yang memutuskan untuk lanjut menggunakan produk. Sedangkan 495,987 konsumen *High Price*, 407,814 konsumen *Medium Price*, dan 703,832 konsumen *High Price* lainnya merespon pembaharuan kontrak dengan tidak lanjut menggunakan produk antivirus 'X'.

Gambar berikut ini merupakan diagram batang untuk variabel Tipe Konsumen yang terdiri atas konsumen perseorangan dan konsumen organisasi/perusahaan. Dari gambar tersebut diketahui bahwa mayoritas konsumen Perusahaan ‘X’ adalah berbentuk perseorangan. Hal tersebut ditunjukkan oleh adanya 476459 konsumen pada produk *Low Price*, 376825 konsumen produk *Medium Price*, dan 670100 konsumen produk *High Price* yang berbentuk individu. Sedangkan konsumen yang berbentuk organisasi hanya terdapat 23541 orang dari produk *Low Price*, 31985 orang dari produk *Medium Price*, dan 39799 orang dari produk *High Price*.



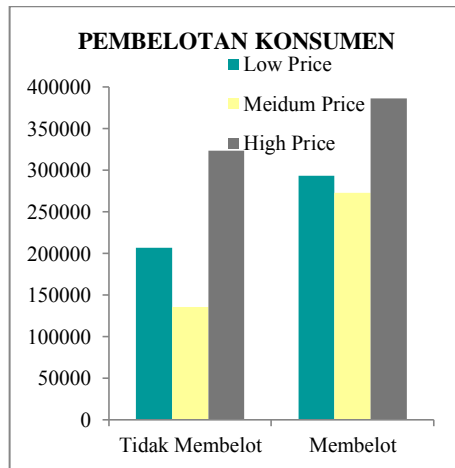
Gambar 4.2 Diagram Batang Tipe Konsumen

Gambar berikut ini adalah diagram batang Status Pengiriman. Dari gambar tersebut, dapat diketahui bahwa lebih banyak *e-mail* yang sukses terkirim daripada yang tidak. Di antara dua kali pengiriman *e-mail* pembaharuan kontrak yang dikirim Perusahaan ‘X’ kepada tiga jenis konsumen, terdapat 300067 *e-mail* yang langsung terkirim ke konsumen *Low Price*, 278279 *e-mail* yang langsung terkirim ke konsumen *Medium Price*, serta 461658 *e-mail* yang langsung terkirim ke konsumen *High Price*. Sedangkan terdapat sisanya gagal tersampaikan ke konsumen lainnya.



Gambar 4.3 Diagram Batang Status Pengiriman

Berikut ini merupakan diagram batang variabel Pembelotan Konsumen.



Gambar 4.4 Diagram Batang Pembelotan Konsumen

Diagram batang di atas memperlihatkan bahwa terdapat lebih banyak konsumen Perusahaan 'X' yang tidak membelot. Kon-

sumen *Low Price* yang membelot terdapat sebanyak 293102 orang, konsumen *Medium Price* yang membelot ada sebanyak 273083 orang, sedangkan konsumen *High Price* yang membelot ada sebanyak 386431 orang. Terdapat 206898 konsumen *Low Price*, 135727 konsumen *Medium Price*, dan 323468 konsumen *High Price* yang tidak membelot.

4.2 Analisis Binary Logistic Regression

Analisis *Binary Logistic Regression* dilakukan untuk mendapatkan model yang dapat menggambarkan hubungan antara Pembelotan Konsumen dengan Akumulasi Update, Harga Produk, Jawaban Kontrak, Tipe Konsumen, serta Status Pengiriman. Analisis ini akan dilakukan pada data *Low Price*, *Medium Price*, dan *High Price* dengan metode evaluasi *Holdout*. Dari masing-masing data tersebut, 90% bagian dari data akan digunakan untuk menyusun model, dan 10% bagian data lainnya digunakan untuk menguji model.

Dalam penelitian ini, koefisien parameter model diestimasi dengan menggunakan *Maximum Likelihood Estimation* dari 90% data *training*. Berikut ini merupakan hasil estimasi parameter model regresi logistik biner untuk data *training Low Price*.

Tabel 4.2.1 Koefisien Parameter Model Awal untuk Konsumen *Low Price*

Parameter	Koefisien	Z	P-Value	Keterangan
(Intercept)	1,53	128,81	0	Signifikan
Akumulasi Update	-0,43	-193,99	0	Signifikan
Harga Produk	$2,03 \times 10^{-6}$	1,37	0,17	Tidak Signifikan
Jawaban Kontrak	-2,92	-46,63	0	Signifikan
Tipe Konsumen	0,0057	0,37	0,715	Tidak Signifikan
Status Pengiriman	-0,85	-121,43	0	Signifikan

Dari tabel di atas, dengan tingkat kesalahan 5% dan taraf keyakinan 95% diketahui bahwa variabel Harga Produk dan Tipe Konsumen tidak memberikan pengaruh yang signifikan terhadap model sedangkan variabel lainnya berpengaruh signifikan pada model. Maka model baru dibentuk dengan melibatkan hanya

prediktor yang berpengaruh signifikan saja. Berikut ini hasil estimasi parameter model regresi logistik biner untuk data *training Low Price* tanpa variabel Harga Produk dan Tipe Konsumen.

Tabel 4.2.2 Koefisien Parameter Model Terbaik untuk Konsumen *Low Price*

Parameter	Koefisien	Z	P-Value	Keterangan
(Intercept)	1,54	244,5	$<2 \times e^{-16}$	Signifikan
Akumulasi Update	-0,43	-194,28	$<2 \times e^{-16}$	Signifikan
Jawaban Kontrak)	-2,92	-46,63	$<2 \times e^{-16}$	Signifikan
Status Pengiriman	-0,85	-121,62	$<2 \times e^{-16}$	Signifikan

Dari tabel di atas, diketahui bahwa semua variabel telah berpengaruh signifikan terhadap model. Untuk dapat menginterpretasikan koefisien model, maka nilai *odds ratio* dihitung dengan mencari eksponensial dari nilai-nilai koefisien tersebut.

Tabel 4.2.3 *Odss Ratio* Koefisien Parameter Model Terbaik untuk Konsumen *Low Price*

Parameter	Koefisien	<i>Odds Ratio</i>
(Intercept)	1,54	4,68
Akumulasi Update	-0,43	0,65
Jawaban Kontrak)	-2,92	0,05
Status Pengiriman	-0,85	0,42

Dari nilai *odds ratio* di atas diketahui bahwa konsumen yang melakukan *update* cenderung membelot 0.65 kali daripada konsumen yang tidak melakukan *update* atau 1,55 kali lebih besar untuk tidak membelot daripada konsumen yang tidak melakukan *update*. Prinsip interpretasi *odds ratio* yang sama berlaku untuk semua nilai *odds ratio* yang bernilai kurang dari 1. Konsumen yang mengkonfirmasi lanjut menggunakan suatu produk cenderung tidak membelot 18,63 kali daripada konsumen yang mengkonfirmasi berhenti menggunakan produk Sedangkan konsumen yang berhasil menerima *e-mail* konfirmasi pembaharuan tepat waktu cenderung tidak membelot 2,35 kali daripada konsumen yang gagal menerima *e-mail* konfirmasi pembaharuan pro-

duk. Model regresi yang dapat dibentuk untuk data *Low Price* adalah sebagai berikut.

$$\pi(x) = \frac{e^{1,54-0,43X_1-2,92X_3-0,85X_5}}{1 + e^{1,54-0,43X_1-2,92X_3-0,85X_5}}$$

Setelah itu, analisis dilanjutkan pada konsumen *Medium Price*. Sama seperti analisis produk *Low Price*, pertama-tama model akan dibentuk dari 90% data *training* melalui estimasi parameter dengan *Maximum Likelihood Estimation*. Berikut ini merupakan hasil estimasi parameter untuk data *training Medium Price*.

Tabel 4.2.4 Koefisien Parameter Model Terbaik untuk Konsumen *Medium Price*

Parameter	Koefisien	Z	P-Value	Keterangan
(Intercept)	2,9	182,608	$<2 \times e^{-16}$	Signifikan
Akumulasi Update	-0,7	-262,973	$<2 \times e^{-16}$	Signifikan
Harga Produk	-0,00011	-51,708	$<2 \times e^{-16}$	Signifikan
Jawaban Kontrak	-3,33	-28,804	$<2 \times e^{-16}$	Signifikan
Tipe Konsumen	-0,053	-3,471	0,000518	Signifikan
Status Pengiriman	-0,2	-22,003	$<2 \times e^{-16}$	Signifikan

Berdasarkan informasi pada tabel di atas, dapat diketahui bahwa semua variabel telah berpengaruh signifikan terhadap model, karena semua nilai *P-Value* untuk semua koefisien telah lebih kecil dari tingkat keasalahan 5%. Dalam *Binary Logistic Regression*, interpretasi koefisien-koefisien tersebut didapatkan dari nilai *odds ratio*. Nilai *odds ratio* dihitung dengan mencari eksponensial dari nilai-nilai koefisien tersebut dan hasilnya adalah sebagai berikut.

Tabel 4.2.5 *Odss Ratio* Koefisien Parameter Model Terbaik Konsumen *Medium Price*

Parameter	Koefisien	Odds Ratio
(Intercept)	2,9	18,24
Akumulasi Update	-0,7	0,49
Harga Produk	-0,00011	0,99
Jawaban Kontrak	-3,33	0,04
Tipe Konsumen	-0,053	0,95
Status Pengiriman	-0,2	0,82

Konsumen yang melakukan *update* dengan melakukan pembelian atau pembaharuan memiliki kecenderungan untuk tidak membelot 2,02 kali lebih besar daripada konsumen yang tidak melakukan *update*. Nilai *odds ratio* Jawaban Kontrak menunjukkan bahwa konsumen yang mengkonfirmasi lanjut menggunakan suatu produk cenderung tidak membelot 25,44 kali daripada konsumen yang mengkonfirmasi berhenti menggunakan produk. Sedangkan konsumen yang berhasil menerima *e-mail* konfirmasi pembaharuan tepat waktu cenderung tidak membelot 1,22 kali daripada konsumen yang gagal menerima *e-mail* konfirmasi pembaharuan produk. Semakin mahalnya harga produk meningkatkan kecenderungan konsumen untuk membelot 1 kali lebih besar daripada yang tidak. Hal tersebut ditunjukkan oleh *odds ratio* Harga Produk yang bernilai 0,99. Nilai tersebut mendekati nilai 1 yang tidak memberikan makna apapun tentang kecenderungan konsumen untuk membelot berdasarkan banyak uang yang ia keluarkan terakhir kali untuk membeli produk *Medium Price*. Begitu pula *odds ratio* Tipe Konsumen menunjukkan bahwa konsumen yang berbentuk organisasi memiliki kecenderungan untuk membelot 1,05 kali lebih besar daripada konsumen yang berbentuk perseorangan. Seperti pada analisis sebelumnya, model yang dibentuk dari analisis *Binary Logistic Regression* adalah berbentuk persamaan logit. Model logit regresi yang dapat dibentuk untuk data *Medium Price* adalah sebagai berikut.

$$y(x) = \frac{e^{2,9-0,705X_1-0,00011X_2-3,326X_3-0,0528X_4-0,201X_5}}{1 + e^{2,9-0,705X_1-0,00011X_2-3,326X_3-0,0528X_4-0,201X_5}}$$

Pada hasil analisis di atas terdapat dua nilai parameter yang mendekati nol dan nilai *odds ratio* yang mendekati 1 untuk suatu variabel Harga Produk dan Tipe Konsumen, namun variabel tersebut teruji berpengaruh signifikan terhadap model. Hal ini merupakan kelemahan *Binary Logistic Regression* dalam menangani *Big Data*. Telah disebutkan sebelumnya bahwa dalam kasus *Big Data*, *Binary Logistic Regression* cenderung menghasilkan keputusan uji variabel yang berpengaruh signifikan, karena nilai *P-Value*

lue pada uji signifikansi cenderung jatuh mendekati nol. Hal tersebut mengakibatkan hasil analisis dengan *Binary Logistic Regression* kurang dapat dipercaya. Mengingat sebagai pendekatan parametrik, *Binary Logistic Regression* juga melibatkan pengujian parameter.

Analisis *Binary Logistic Regression* terakhir dilakukan untuk konsumen *High Price*. Berikut ini merupakan hasil estimasi koefisien-koefisien tersebut untuk data *High Price*.

Tabel 4.2.6 Koefisien Parameter Model Terbaik untuk Konsumen *High Price*

Parameter	Koefisien	Z	P-Value	Keterangan
(Intercept)	1,68	167,27	$<2 \times e^{-16}$	Signifikan
Akumulasi Update	-0,54	-274,11	$<2 \times e^{-16}$	Signifikan
Harga Produk	-0,000008	-9,07	$<2 \times e^{-16}$	Signifikan
Jawaban Kontrak	-2,84	-55,21	$<2 \times e^{-16}$	Signifikan
Tipe Konsumen	-0,034	-2,84	0.00446	Signifikan
Status Pengiriman	-0,92	-148,77	$<2 \times e^{-16}$	Signifikan

P-Value masing-masing koefisien dari analisis *Binary Logistic Regression* di atas bernilai lebih kecil dari batas kesalahan yang ditetapkan 5%. Sehingga dapat dinyatakan bahwa semua parameter di atas berpengaruh secara signifikan pada model. Guna menginterpretasikan nilai-nilai tersebut, maka masing-masing nilai koefisien dieksponensialkan dan nilai *odds ratio* didapatkan sebagai berikut. Nilai-nilai ini menjelaskan kecenderungan konsumen untuk membelot berdasarkan variabel-variabel yang mempengaruhinya.

Tabel 4.2.7 *Odss Ratio* Koefisien Parameter Model Terbaik untuk Konsumen *High Price*

Parameter	Koefisien	<i>Odds Ratio</i>
(Intercept)	1,68	5,37
Akumulasi Update	-0,5385	0,58
Harga Produk	-0,00000797	0,99
Jawaban Kontrak	-2,845	0,06
Tipe Konsumen	-0,03451	0,97
Status Pengiriman	-0,9198	0,39

Dari nilai *odds ratio* di atas diketahui bahwa konsumen *High Price* yang melakukan *update* dari pembelian atau pembaharuan memiliki kecenderungan untuk tidak membelot 1/0,8 atau 1,71 kali lebih besar daripada konsumen yang tidak melakukan *update*. Interpretasi *odds ratio* untuk variabel Harga Produk pada data *High Price* juga sama seperti yang terjadi pada data *Medium Price*. Variabel harga tersebut tidak menunjukkan kecenderungan pembelotan konsumen yang berarti. Hal tersebut dikarenakan nilai *odds ratio* bernilai mendekati 1. Sedangkan konsumen *High Price* yang mengkonfirmasi lanjut menggunakan suatu produk cenderung tidak membelot 1/0,06 atau 17,21 kali daripada konsumen yang mengkonfirmasi berhenti menggunakan produk. Hal yang terjadi di variabel Harga Produk, terjadi pula di variabel Tipe Konsumen. Seperti yang terjadi pada data *High Price*, variabel ini juga tidak memberikan arti kecenderungan konsumen berdasarkan jenis konsumen tersebut. Sedangkan konsumen *High Price* yang berhasil menerima *e-mail* konfirmasi pembaharuan tepat waktu cenderung tidak membelot 2,51 kali daripada konsumen yang gagal menerima *e-mail* konfirmasi pembaharuan produk. Berikut ini adalah model regresi yang dapat dibentuk untuk konsumen *High Price*.

$$y(x) = \frac{e^{1,68-0,5385X_1-0,00000797X_2-2,845X_3-0,03451X_4-0,9198X_5}}{1 + e^{1,68-0,5385X_1-0,00000797X_2-2,845X_3-0,03451X_4-0,9198X_5}}$$

Meskipun model *Binary Logistic Regression* pada konsumen *Medium Price* dan *High Price* yang didapat tidak memuaskan, analisis akan dilanjutkan sesuai dengan tujuan penelitian. Nilai variabel prediktor pada data *testing* masing-masing konsumen disubstitusikan ke dalam model di atas dan nilai probabilitaspun di dapatkan. Konsumen diklasifikan dengan membandingkan nilai probabilitas tiap konsumen dengan *threshold* 0,5. Konsumen dengan nilai probabilitas lebih dari 0,5 akan diklasifikasikan ke dalam kelas membelot, sedangkan konsumen-konsumen dengan probabilitas kurang dari 0,5 diputuskan tidak membelot. Berikut ini merupakan tabulasi silang hasil klasifikasi ketiga konsumen.

Tabel 4.2.8 Tabulasi Silang Klasifikasi dari *Binary Logistic Regression*

		Kelas Aktual					
		Low Price		Medium Price		High Price	
		<i>p</i> (+)	<i>n</i> (-)	<i>p</i> (+)	<i>n</i> (-)	<i>p</i> (+)	<i>n</i> (-)
Kelas Prediksi	<i>p</i> (+)	22878	10299	24467	6432	30337	14113
	<i>n</i> (-)	6433	10391	2842	7141	8307	18234

Kelas positif (*p*) pada tabel di atas mewakili kelas membelot sedangkan kelas negatif (*n*) mewakili kelas tidak membelot. Dari tabel di atas didapatkan bahwa terdapat 22878 konsumen *Low Price* yang sebenarnya memang membelot dan akhirnya diprediksi membelot, sedangkan konsumen yang sebenarnya membelot namun diprediksi tidak membelot adalah sebanyak 6433 orang. Terdapat 10391 konsumen tidak membelot yang diprediksi tidak membelot, sedangkan 10299 konsumen sisanya diprediksi membelot. Cara interpretasi yang sama diberlakukan untuk hasil klasifikasi konsumen untuk produk lainnya. Berikut ini merupakan ketepatan klasifikasi yang didapatkan dari hasil klasifikasi di atas.

Tabel 4.2.9 Ketepatan Klasifikasi Analisis *Binary Logistic Regression*

Produk	Akurasi	<i>Sensitivity</i>	<i>Spesificity</i>
<i>Low Price</i>	66,54	77,31	68,42
<i>Medium Price</i>	78,05	89,6	78,5
<i>High Price</i>	50,22	52,61	56,37

Binary Logistic Regression telah mampu menebak 66,54% dari konsumen *Low Price* dengan benar, 77,31% konsumen *Low Price* yang membelot dengan benar dan 68,42% konsumen *Low Price* yang tidak membelot dengan benar. Cara interpretasi yang sama berlaku pada konsumen lainnya. Dari nilai di atas, berikut ini adalah ukuran kelasahan klasifikasi yang didapat dari *Binary Logistic Regression*.

Tabel 4.2.10 Kesalahan Klasifikasi Analisis *Binary Logistic Regression*

Produk	Error	FNR	FPR
<i>Low Price</i>	33.46	22.69	31.58
<i>Medium Price</i>	21.95	10.4	21.5
<i>High Price</i>	49.78	47.39	43.63

4.3 Analisis LORENS

Analisis dengan metode *Binary Logistic Regression* terkendala hasil uji signifikansi parameter yang cenderung menyimpulkan bahwa parameter yang diuji signifikan. Hal tersebut terbukti dengan diperolehnya nilai *P-Value* yang sangat kecil pada data *Medium Price* dan *High Price*, tepatnya pada variabel Harga Produk dan Tipe Konsumen. Kedua variabel tersebut berpengaruh signifikan terhadap model namun memiliki nilai koefisien yang sangat kecil. Sehingga nilai *odds ratio* kedua variabel tersebut tidak menceritakan kecenderungan pembelotan apapun. Hal ini mengakibatkan hasil klasifikasi dengan *Binary Logistic Regression* untuk ketiga data *Low Price*, *Medium Price*, dan *High Price* kurang dapat dipercaya. Maka dari itu, selanjutnya ketiga data tersebut dianalisis dengan metode LORENS yang merupakan pendekatan klasifikasi secara komputasional dan tidak melibatkan pengujian parameter. Dengan menggunakan data *training* dan data *testing* yang sama dengan yang digunakan dalam analisis *Binary Logistic Regression*, metode evaluasi *Holdout* dilakukan sebagai pembanding yang adil terhadap hasil analisis dengan *Binary Logistic Regression*.

Sebagai pendekatan komputasional dalam menyelesaikan kasus klasifikasi, LORENS tidak memiliki asumsi apapun untuk dipenuhi. Demi mencapai hasil klasifikasi terbaik, analisis LORENS untuk satu jenis produk akan dilakukan beberapa kali dengan besar beberapa partisi yang berbeda yang mungkin dibentuk. Dalam kasus ini partisi yang mungkin dibentuk adalah sebesar 1, 2, 3, 4, dan 5 partisi. LORENS dapat menemukan nilai *trheshold* optimum, namun analisis LORENS dengan *threshold* 0,5 akan tetap dilakukan sebagai perbandingan. Sedangkan nilai *ensemble* untuk analisis ini ditetapkan sebesar 10 karena menurut penelitian sebelumnya relatif dapat menghasilkan hasil akurasi klasifikasi yang baik.

Dengan ukuran *ensemble* 10, banyak model yang dibentuk pada analisis LORENS dengan *Holdout* untuk 1 partisi adalah 10 model, untuk 2 partisi adalah 20 model, untuk 3 partisi adalah 30

model, untuk 4 partisi adalah 40 model, dan dengan 5 partisi adalah 50 model.

Untuk menggunakan metode evaluasi *Holdout*, 10% bagian dari data diambil secara acak sebagai data *testing* dengan proporsi kelas respon yang representatif terhadap data keseluruhan. Kemudian 90% data digunakan untuk menyusun model. Data *training* dan *testing* dalam analisis ini sama dengan yang telah digunakan dalam metode *Binary Logistic Regression*.

Dalam analisis LORENS, prediktor-prediktor dialokasikan ke dalam ruang-ruang partisi untuk kemudian dibentuk model pada masing-masing subruang partisi tersebut. Proses tersebut akan diulang berkali-kali sebanyak ukuran *ensemble* yang ditetapkan. Misalkan saja analisis LORENS akan dilakukan terhadap data *Low Price* dengan parameter ukuran *threshold*, ukuran partisi, dan ukuran *ensemble*. Ukuran partisi yang ditentukan adalah 3 partisi dan ukuran *ensemble* yang ditentukan adalah sebanyak 10 *ensemble*. Maka 5 variabel prediktor dalam penelitian ini akan dialokasikan ke dalam 3 ruang partisi dan proses tersebut akan diulang sampai sebanyak 10 kali dari jumlah *ensemble* yang ditentukan.

Tabel 4.3.1 *Random Sampling* Variabel Prediktor pada Ruang Partisi

Variabel Prediktor	<i>Ensemble</i>									
	1	2	3	4	5	6	7	8	9	10
Akumulasi Update	3	2	3	1	1	1	2	1	1	3
Harga Produk	1	3	1	1	3	2	3	1	2	3
Jawaban Kontrak	3	3	3	3	2	3	1	3	1	1
Tipe Konsumen	1	1	1	3	3	1	1	2	3	1
Status Pengiriman	2	1	2	2	1	3	3	3	3	2

Tabel di atas menunjukkan pengalokasian variabel prediktor ke dalam ruang partisi. Pada *ensemble* pertama variabel Harga Produk dan Tipe Konsumen dialokasikan ke dalam ruang partisi pertama, Status Pengiriman dialokasikan ke dalam ruang partisi ke-2, sedangkan Akumulasi Update dan Jawaban Kontrak dialokasikan ke dalam ruang partisi ke-3, cara mengalokasikan yang sama dilakukan pada *ensem-ble* ke-2 hingga ke-10.

Tabel 4.3.2 Model Regresi Logistik pada Ruang Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3
1	$e^{0,2+2,6 \times 10^{-5}X_2-0,02X_4}$	$e^{1,05-1,1X_5}$	$e^{1,08-0,5X_1-3,05X_3}$
2	$1 + e^{0,2+2,6 \times 10^{-5}X_2-0,02X_4}$	$1 + e^{1,05-1,1X_5}$	$1 + e^{1,08-0,5X_1-3,05X_3}$
3	$e^{1,05-0,02X_4-1,1X_5}$	$e^{1,05-0,5X_1}$	$e^{0,2+2,7 \times 10^{-5}X_2-2,8X_3}$
4	$1 + e^{1,05-0,02X_4-1,1X_5}$	$1 + e^{1,05-0,5X_1}$	$1 + e^{0,2+2,7 \times 10^{-5}X_2-2,8X_3}$
5	$e^{0,2+2,6 \times 10^{-5}X_2-0,02X_4}$	$e^{1,05-1,1X_5}$	$e^{1,08-0,5X_1-3,05X_3}$
6	$1 + e^{0,2+2,6 \times 10^{-5}X_2-0,02X_4}$	$1 + e^{1,05-1,1X_5}$	$1 + e^{1,08-0,5X_1-3,05X_3}$
7	$e^{1-0,5X_1+9 \times 10^{-6}X_2}$	$e^{1,05-1,1X_5}$	$e^{0,4-2,8X_3-0,01X_4}$
8	$1 + e^{1-0,5X_1+9 \times 10^{-6}X_2}$	$1 + e^{1,05-1,1X_5}$	$1 + e^{0,4-2,8X_3-0,01X_4}$
9	$e^{1,5-0,4X_1-0,9X_5}$	$e^{0,4-2,8X_3}$	$e^{0,2+2,6 \times 10^{-5}X_2-0,02X_4}$
10	$1 + e^{1,5-0,4X_1-0,9X_5}$	$1 + e^{0,4-2,8X_3}$	$1 + e^{0,2+2,6 \times 10^{-5}X_2-0,02X_4}$
	$e^{1,05-0,5X_1+0,03X_4}$	$e^{0,2+2,6 \times 10^{-5}X_2}$	$e^{1,05-0,02X_4-1,1X_5}$
	$1 + e^{1,05-0,5X_1+0,03X_4}$	$1 + e^{0,2+2,6 \times 10^{-5}X_2}$	$1 + e^{1,05-0,02X_4-1,1X_5}$
	$e^{0,4-2,8X_3-0,01X_4}$	$e^{1,05-0,5X_1}$	$e^{1+1,4 \times 10^{-5}X_2X_2-1,1X_5}$
	$1 + e^{0,4-2,8X_3-0,01X_4}$	$1 + e^{1,05-0,5X_1}$	$1 + e^{1+1,4 \times 10^{-5}X_2X_2-1,1X_5}$
	$e^{1-0,5X_1+9 \times 10^{-6}X_2}$	$e^{0,3-0,008X_4}$	$e^{1,06-2,7X_3-1,1X_5}$
	$1 + e^{1-0,5X_1+9 \times 10^{-6}X_2}$	$1 + e^{0,3-0,008X_4}$	$1 + e^{1,06-2,7X_3-1,1X_5}$
	$e^{1,08-0,5X_1-3,05X_3}$	$e^{0,2+0,00003X_2}$	$e^{1,05-0,02X_4-1,1X_5}$
	$1 + e^{1,08-0,5X_1-3,05X_3}$	$1 + e^{0,2+0,00003X_2}$	$1 + e^{1,05-0,02X_4-1,1X_5}$
	$e^{0,4-2,8X_3-0,01X_4}$	$e^{1,05-1,1X_5}$	$e^{1-0,5X_1+9 \times 10^{-6}X_2}$
	$1 + e^{0,4-2,8X_3-0,01X_4}$	$1 + e^{1,05-1,1X_5}$	$1 + e^{1-0,5X_1+9 \times 10^{-6}X_2}$

Tabel di atas menunjukkan model regresi logistik yang terbentuk pada masing-masing ruang partisi dan pada 10 *ensemble*. Pengambilan variabel untuk dialokasikan dalam LORENS dilakukan dengan menggunakan random *sampling*. Dengan variabel-variabel yang sudah disebar, model regresi logistik akan dibentuk pada masing-masing ruang partisi dari variabel-variabel tersebut. Tabel di atas menunjukkan bahwa pada *ensemble* pertama, model partisi pertama dibentuk oleh variabel Harga Produk (X_2) dan Tipe Konsumen (X_4), model pada ruang partisi ke-2 disusun oleh variabel Status Pengiriman (X_5), sedangkan model pada ruang partisi ke-3 disusun oleh variabel Akumulasi Update (X_1) dan Jawaban Kontrak (X_3). Cara membaca yang sama berlaku pada *ensemble* ke-2, *ensemble* ke-3 hingga *ensemble* ke-10.

Setelah model didapatkan, maka langkah selanjutnya adalah mengklasifikasikan data *testing* dengan ke dalam model-model yang telah terbentuk. Namun sebelum itu, diperlukan terlebih dahulu nilai *threshold* yang akan menjadi batas keputusan apakah pengamatan dikelompokkan ke dalam kelas membelot atau tidak. Nilai *threshold* yang dapat digunakan adalah 0,5 atau nilai *threshold* optimum yang dapat ditemukan LORENS. Variabel data *testing* yang disubstitusikan ke dalam model akan menghasilkan nilai-nilai probabilitas. Nilai probabilitas akhir untuk masing-masing konsumen didapatkan dari rata-rata nilai probabilitas konsumen tersebut dari ketiga model pada tiga ruang partisi di atas. Apabila nilai probabilitas akhir tersebut lebih besar dari nilai *threshold* optimum, maka pengamatan/ konsumen akan dikategorikan ke dalam kelas 1 yang merupakan kelas konsumen membelot. Namun apabila sebaliknya, maka konsumen akan dikategorikan ke dalam kelas tidak membelot. Nilai *threshold* optimum LORENS dalam analisis ini didapatkan dari menjumlahkan nilai rata-rata variabel respon data *training* dengan nilai *threshold* umum yaitu 0,5 kemudian membagi keduanya dengan 2. Dalam analisis LORENS untuk data *Low Price* dengan metode evaluasi *Holdout* ini, nilai *threshold* yang didapatkan adalah sebesar 0,5431. Dari nilai *threshold* optimum inilah konsumen diklasifikasikan menjadi membelot atau tidak.

Proses di atas diulangi untuk *ensemble* ke-2, ke-3, hingga *ensemble* ke-10. Pada tahap ini telah dimiliki 10 hasil klasifikasi pada 10 *ensemble*. Hasil klasifikasi akhir yang digunakan didapatkan dari memilih kelas klasifikasi mayoritas masing-masing pengamatan di seluruh *ensemble* untuk pengamatan dengan indeks yang sama. Apabila 6 atau lebih hasil klasifikasi dari 10 *ensemble* menyatakan bahwa seorang konsumen dikategorikan sebagai membelot, maka keputusan akhir konsumen tersebut dikategorikan membelot dan apabila terdapat lebih dari 5 hasil klasifikasi di antara 10 *ensemble* yang menyatakan seorang konsumen tidak membelot, maka konsumen dikategorikan tidak membelot.

Akhirnya hasil klasifikasi data *Low Price* dengan ukuran partisi 3 dapat ditampilkan dalam tabulasi silang seperti berikut.

Tabel 4.3.3 Tabulasi Silang Klasifikasi Konsumen *Low Price* untuk 3 Partisi dan *Threshold* Optimum

		Kelas Aktual	
		<i>p</i> (+)	<i>n</i> (–)
Kelas Prediksi	<i>p</i> (+)	22932	10497
	<i>n</i> (–)	6379	10193

Berdasarkan tabel tersebut, dapat diketahui bahwa terdapat 22932 konsumen yang benar-benar membelot dan berakhir diprediksi membelot, sedangkan konsumen yang membelot namun diprediksi tidak membelot ada sebanyak 6379 orang. Dalam memprediksi konsumen tidak membelot, terdapat 10193 konsumen yang sebenarnya tidak membelot dan akhirnya diprediksi tidak membelot, sedangkan 10297 konsumen sisanya diprediksi membelot.

Ukuran akurasi hasil di atas adalah 0,662487 dengan *sensitivity* 0,782368, dan *specificity* 0,492653. LORENS dengan 3 partisi telah mampu menebak 66,24% keputusan konsumen dengan tepat dengan tingkat ketepatan menebak konsumen yang membelot sebesar 78,23% dan ketepatan menebak konsumen yang tidak membelot sebesar 49,26%. Ukuran kesalahan klasifikasi hasil analisis di atas adalah *Error* sebesar 33,76%, FNR 21, 77%, dan FPR sebesar 50,74%. Sehingga LORENS dengan 3 partisi tidak mampu menebak 33,67% keputusan konsumen dengan tepat dengan tingkat kesalahan menebak konsumen membelot menjadi tidak membelot sebesar 21,77% dan konsumen tidak membelot menjadi membelot sebesar 50,74%.

Dengan prosedur yang sama analisis LORENS mulai dari pembagian data hingga pengklasifikasian dan penghitungan ketepatan klasifikasi diterapkan pada data *Low Price*, *Medium Price*, dan *High Price*. Masing-masing dengan ukuran partisi 1 hingga 5, baik dengan *threshold* 0,5 maupun *threshold* optimum. Berikut ini merupakan hasil klasifikasi untuk data *Low Price* untuk masing-masing ukuran partisi.

Tabel 4.3.4 Tabulasi Silang Klasifikasi Hasil Analisis LORENS untuk Konsumen *Low Price*

		Kelas Aktual			
		Threshold Optimum		Threshold 0,5	
		$p(+)$	$n(-)$	$p(+)$	$n(-)$
Kelas Prediksi	1 Partisi	$p(+)$	22878	10299	22878
		$n(-)$	6433	10391	6433
	2 Partisi	$p(+)$	22878	10299	25821
		$n(-)$	6433	10391	3490
	3 Partisi	$p(+)$	22932	10497	26247
		$n(-)$	6379	10193	3064
	4 Partisi	$p(+)$	25920	14089	27873
		$n(-)$	3391	6601	1438
	5 Partisi	$p(+)$	25937	14125	28919
		$n(-)$	3374	6565	392

Berikut ini merupakan tabulasi silang yang memuat hasil klasifikasi LORENS untuk data *Medium Price*, masing-masing dengan ukuran partisi 1 sampai 5, serta ukuran *threshold* optimum dan *threshold* 0,5..

Tabel 4.3.5 Tabulasi Silang Klasifikasi Hasil Analisis LORENS untuk Konsumen *Medium Price*

		Kelas Aktual			
		Threshold Optimum		Threshold 0,5	
		$p(+)$	$n(-)$	$p(+)$	$n(-)$
Kelas Prediksi	1 Partisi	$p(+)$	21519	4246	24467
		$n(-)$	5790	9327	2842
	2 Partisi	$p(+)$	24057	6068	26335
		$n(-)$	3252	7505	974
	3 Partisi	$p(+)$	25148	8252	27224
		$n(-)$	2161	5321	85
	4 Partisi	$p(+)$	25211	8505	27306
		$n(-)$	2098	5068	3
	5 Partisi	$p(+)$	26753	12611	27307
		$n(-)$	556	962	2

Sedangkan berikut ini merupakan hasil klasifikasi untuk data *Medium Price* untuk masing-masing ukuran partisi dengan *threshold* 0,5 dan optimum.

Tabel 4.3.6 Tabulasi Silang Klasifikasi Hasil Analisis LORENS untuk Konsumen *High Price*

		Kelas Aktual			
		Threshold Optimum		Threshold 0,5	
		$p(+)$	$n(-)$	$p(+)$	$n(-)$
Kelas Prediksi	1	$p(+)$	29747	13081	30337
	Partisi	$n(-)$	8897	19266	8307
	2	$p(+)$	30471	14631	30513
	Partisi	$n(-)$	8173	17716	8131
	3	$p(+)$	30513	14740	30520
	Partisi	$n(-)$	8131	17607	8124
	4	$p(+)$	30519	14740	34171
	Partisi	$n(-)$	8125	17607	4473
	5	$p(+)$	30521	14783	34629
	Partisi	$n(-)$	8123	17564	4015

Dari hasil klasifikasi di atas, kemudian dicari ukuran ketepatan klasifikasi yaitu akurasi, *sensitivity* dan *specificity*. Akurasi merupakan ketepatan dalam menebak konsumen membelot dan tidak membelot dengan benar. *Sensitivity* merupakan ketepatan dalam menebak konsumen membelot di antara konsumen-konsumen yang benar-benar membelot. Sedangkan *Specificity* merupakan ketepatan menebak konsumen tidak membelot di antara konsumen yang benar-benar tidak membelot.

Nilai akurasi dan *specificity* pada semua produk mengecil seiring semakin besarnya partisi. Sedangkan nilai *sensitivity* semakin membesar seiring besarnya ukuran partisi. *Threshold* optimum menjaga perubahan ukuran ketepatan klasifikasi menjadi lebih kecil daripada yang bisa dihasilkan dengan *threshold* 0,5. Dalam kasus ini kelemahan metode LORENS ditunjukkan dari kecenderungan berubahnya nilai ketepatan klasifikasi seiring berubahnya partisi. Hal ini disebabkan oleh ketidak-mampuan LORENS untuk menghitung nilai *threshold* optimum jika jumlah prediktor lebih kecil dari jumlah pengamatan. Nilai *threshold* yang didapatkan dari perhitungan cenderung mengkategorikan pengamatan menjadi membelot, padahal proporsi kelas pengamatan data dalam kasus ini tidak seimbang. Berikut ini merupa-

kan nilai akurasi, *sensitivity*, dan *specificity* dari analisis LOR-ENS data ketiga produk dengan metode evaluasi *Holdout*.

Tabel 4.3.7 Ukuran Ketepatan Klasifikasi Hasil Analisis LORENS

		<i>Threshold Optimum</i>				
Produk	Partisi	1	2	3	4	5
Low Price	Akurasi	66,54	66,54	66,25	65,04	65,00
	<i>Sensitivity</i>	78,05	78,05	78,24	88,43	88,49
	<i>Specificity</i>	50,22	50,22	49,27	31,90	31,73
Medium Price	Akurasi	75,45	77,20	74,53	74,06	67,79
	<i>Sensitivity</i>	78,80	88,09	92,09	92,32	97,96
	<i>Specificity</i>	68,72	55,29	39,20	37,34	7,09
High Price	Akurasi	69,04	67,88	67,78	67,79	67,73
	<i>Sensitivity</i>	76,98	78,85	78,96	78,97	78,98
	<i>Specificity</i>	59,56	54,77	54,43	54,43	54,30
		<i>Threshold 0,5</i>				
Produk	Partisi	1	2	3	4	5
Low Price	Akurasi	66,54	65,04	64,50	63,41	59,27
	<i>Sensitivity</i>	78,05	88,09	89,55	95,09	98,66
	<i>Specificity</i>	50,22	32,39	29,01	18,54	3,47
Medium Price	Akurasi	77,32	69,60	66,75	66,87	66,83
	<i>Sensitivity</i>	89,59	96,43	99,69	99,99	99,99
	<i>Specificity</i>	52,61	15,62	0,49	0,24	0,10
High Price	Akurasi	68,42	67,78	67,74	65,58	65,95
	<i>Sensitivity</i>	78,50	78,96	78,98	88,43	89,61
	<i>Specificity</i>	56,37	54,43	54,32	38,29	37,68

Penurunan akurasi dan *specificity* serta kenaikan *sensitivity* lebih terjaga dengan *threshold* optimum. Dalam analisis LOR-ENS ini pemilihan ukuran partisi terbaik tidak dilandaskan pada partisi yang memberikan ketepatan akurasi terbaik karena analisis LORENS untuk data konsumen dalam kasus ini memberikan hasil terbaik untuk ukuran 1 partisi, yang artinya tidak ada partisi. Model yang dikehendaki dalam LORENS adalah model yang parsimoni. Penambahan jumlah variabel dalam model dilakukan secara bertahap dari penentuan ukuran partisi 5 saat masing-masing model hanya mengandung 1 variabel hingga partisi 1 saat semua variabel termasuk dalam model. Peningkatan ketepatan klasifikasi

akibat dari bertambahnya variabel dalam modelpun diamati. Ukuran partisi terbaik ditentukan saat penambahan variabel ke dalam model dapat meningkatkan ketepatan klasifikasi dengan cukup baik pertama kali, sedangkan penambahan variabel di tingkat berikutnya tidak meningkatkan ketepatan klasifikasi tidak lebih baik dari peningkatan sebelumnya.

Berdasarkan tabel di atas, dapat dilihat bahwa dengan *threshold* optimum, LORENS mampu menghasilkan ketepatan klasifikasi yang lebih baik dan lebih seimbang, baik untuk kelas konsumen membelot maupun tidak membelot. Berdasarkan hasil tersebut, maka hasil analisis LORENS dengan *threshold* optimum yang akan digunakan, sehingga dalam tahap berikutnya hasil analisis LORENS dengan *threshold* 0,5 tidak dipertimbangkan. Dari sinilah pencarian partisi optimum dimulai. Berikut ini merupakan tabel yang memperlihatkan ukuran partisi optimum dari analisis LORENS dengan *Holdout* untuk masing-masing produk. Nilai perubahan ketepatan di bawah ini merupakan selisih dari nilai ketepatan klasifikasi pada suatu ukuran partisi dengan ketepatan klasifikasi pada ukuran partisi yang lebih rendah sebelumnya.

Tabel 4.3.8 Ukuran Partisi Optimum dari Analisis LORENS

Produk	Threshold	Partisi Optimum	Perubahan Nilai		
			Akur.	Sens.	Spec.
<i>Low Price</i>	Optimum	3	1,21	-10,19	17,36
<i>Medium Price</i>	Optimum	4	6,27	-5,65	30,25
<i>High Price</i>	Optimum	1	1,16	-1,87	4,79

Tabel di atas menjelaskan bahwa ukuran partisi optimum untuk analisis LORENS konsumen *Low Price* dengan *threshold* optimum adalah 3 partisi. Pada data *Low Price* terdapat 5 variabel yang dialokasikan ke dalam 3 ruang partisi sehingga terdapat 2 ruang partisi berisi 2 variabel dan 1 partisi sisanya terdiri atas 1 variabel. Ukuran partisi optimum untuk analisis LORENS dengan *Holdout* untuk konsumen *Medium Price* adalah 4 partisi. Ukuran partisi optimum untuk analisis LORENS dengan *Holdout* untuk konsumen *High Price* adalah 1 partisi. Berikut ini merupakan ukuran kesalahan klasifikasi hasil analisis di atas.

Tabel 4.3.9 Ukuran Kesalahan Klasifikasi Hasil Analisis LORENS

		Threshold Optimum				
Produk	Partisi	1	2	3	4	5
Low Price	<i>Error</i>	33.46	33.46	33.75	34.96	35.00
	FNR	21.95	21.95	21.76	11.57	11.51
	FPR	49.78	49.78	50.73	68.10	68.27
Medium Price	<i>Error</i>	24.55	22.80	25.47	25.94	32.21
	FNR	21.20	11.91	7.91	7.68	2.04
	FPR	31.28	44.71	60.80	62.66	92.91
High Price	<i>Error</i>	30.96	32.12	32.22	32.21	32.27
	FNR	23.02	21.15	21.04	21.03	21.02
	FPR	40.44	45.23	45.57	45.57	45.70

		Threshold 0,5				
Produk	Partisi	1	2	3	4	5
Low Price	<i>Error</i>	33.46	34.96	35.50	36.59	40.73
	FNR	21.95	11.91	10.45	4.91	1.34
	FPR	49.78	67.61	70.99	81.46	96.53
Medium Price	<i>Error</i>	22.68	30.40	33.25	33.13	33.17
	FNR	10.41	3.57	0.31	0.01	0.01
	FPR	47.39	84.38	99.51	99.76	99.90
High Price	<i>Error</i>	31.58	32.22	32.26	34.42	34.05
	FNR	21.50	21.04	21.02	11.57	10.39
	FPR	43.63	45.57	45.68	61.71	62.32

4.4 Analisis LORENS dengan Cross Validation

Metode evaluasi yang dianjurkan LORENS adalah *Cross Validation* karena metode ini memperlakukan semua pengamatan secara adil dalam hal kedudukan sebagai *training set* dan *testing set*. Meskipun LORENS dengan *Cross Validation* tidaklah sebanding dengan *Binary Logistic Regression* maupun LORENS dengan *Holdout* karena susunan data *training* dan *testing* yang berbeda, algoritma yang agregatif dalam LORENS dengan metode evaluasi *Cross Validation* diharapkan dapat meningkatkan performa ketepatan klasifikasi LORENS. Sama seperti analisis LORENS dengan metode *Holdout*, analisis dengan *Cross Validation* dilakukan beberapa kali dengan besar beberapa partisi yang berbeda yang mungkin dibentuk yaitu 1, 2, 3, 4, dan 5 partisi. Ukuran *threshold* yang digunakan adalah *threshold* optimum dan

Threshold 0,5. Sedangkan nilai *ensemble* untuk analisis ini ditetapkan sebesar 10. Dalam analisis ini ukuran *folds* yang digunakan dalam metode *Cross Validation* adalah 10 *folds*.

Dalam analisis ini, data dibagi menjadi 10 bagian. Masing-masing bagian akan diperlakukan sebagai data *testing* pada tiap *fold*. Bagian data pertama akan digunakan sebagai data *testing* di *fold* ke-2 sedangkan sisanya akan digunakan sebagai data *training*, bagian data ke-2 akan digunakan sebagai data *testing* di *fold* ke-2 sedangkan sisanya digunakan sebagai data *training*, dan seterusnya. Banyak model yang dapat dibentuk dengan analisis LORENS dengan *Cross Validation* pada tiap *fold* ada sebanyak model pada analisis LORENS dengan *Holdout*. Sehingga dengan ukuran 10 *ensemble* dan 10 *folds*, maka banyak model yang dibentuk dari data *training* dengan 1 partisi adalah 100 model, untuk 2 partisi adalah 200 model, untuk 3 partisi adalah 300 model, untuk 4 partisi adalah 400 model, dan dengan 5 partisi adalah 500 model. Pengalokasian variabel-variabel prediktor untuk membangun model pada masing-masing ruang partisi juga sama, yaitu menggunakan *random sampling*.

Untuk data *High Priced* dengan 2 partisi, model yang akan terbentuk adalah sebanyak 200 model. Di *fold* pertama, variabel prediktor data *testing* disubstitusikan ke dalam model pada masing-masing ruang partisi. Nilai probabilitas seorang konsumen di kedua partisi pada *ensemble* yang sama dirata-rata, kemudian nilai tersebut probabilitas tersebut dibandingkan dengan nilai *threshold* optimum. Dalam analisis LORENS, *Threshold* yang berbeda didapatkan dari *folds* yang berbeda pula.

Tabel 4.4.1 *Threshold* Optimum Konsumen *High Price* untuk 2 Partisi

<i>Fold</i>	<i>Threshold Optimal</i>	<i>Fold</i>	<i>Threshold Optimal</i>
1	0,522120	6	0,522233
2	0,522158	7	0,522117
3	0,522007	8	0,522003
4	0,522338	9	0,522365
5	0,522134	10	0,522257

Pada *fold* pertama, nilai probabilitas pada masing-masing *ensemble* akan dibandingkan dengan 0,522120. Konsumen de-

ngan nilai probabilitas yang lebih besar dari 0,522120 akan dikategorikan sebagai membelot dan sebaliknya. Keputusan membelot atau tidaknya konsumen di-*voting* dari suara terbanyak di antara 10 *ensemble*. Apabila 6 atau lebih hasil klasifikasi dari 10 *ensemble* menyatakan bahwa seorang konsumen adalah membelot, maka akhirnya konsumen tersebut diklasifikasikan sebagai membelot, dan berlaku sebaliknya. Proses di atas dari pensubstitusian data training ke dalam model di tiap ruang partisi, kemudian membandingkan nilai probabilitas terhadap nilai *Threshold* optimum, hingga mendapatkan klasifikasi akhir dilakukan 10 kali hingga 10 *folds* telah terpenuhi. Setelah perbagian data *testing* pada tiap *fold* telah terklasifikasikan, hasil klasifikasi di tiap *fold* disusun kembali menjadi utuh dengan urutan yang tepat dan hasil klasifikasi untuk keseluruhan konsumen dalam data dapat ditabelkan. Berikut ini merupakan tabulasi silang klasifikasi data *High Price* dengan ukuran 2 partisi.

Tabel 4.4.2 Tabulasi Silang Klasifikasi Konsumen *High Price* dengan *Cross Validation* untuk 2 Partisi dan *Threshold* Optimum

		Kelas Aktual	
		<i>p</i> (+)	<i>n</i> (−)
Kelas Prediksi	<i>p</i> (+)	304091	144277
	<i>n</i> (−)	82340	179191

Terdapat 304091 konsumen produk *High Price* yang membelot dan diprediksi membelot, sedangkan konsumen yang membelot namun diprediksi tidak membelot adalah sebanyak 82340 orang, Terdapat 17 9191 konsumen *High Price* tidak membelot yang diprediksi tidak membelot, sedangkan 144277 konsumen sisanya diprediksi membelot. Ukuran akurasi hasil di atas adalah 0,680776 dengan *Sensitivity* 0,7 86922, dan *Specificity* 0,553968. LORENS dengan 2 partisi telah mampu menebak 68,08% keputusan konsumen *High Price* dengan tepat dengan tingkat ketepatan menebak konsumen *High Price* yang membelot sebesar 78,7% dan ketepatan menebak konsumen *High Price* yang tidak membelot sebesar 55,4%. Dengan prosedur yang sama, klasifikasi dapat dilakukan untuk data *Low Price*, *Medium Price*, dan *High*

Price baik dengan *Threshold* 0,5 maupun *Threshold* optimum. Berikut ini merupakan hasil klasifikasi untuk konsumen *Low Price* untuk masing-masing ukuran partisi.

Tabel 4.4.3 Tabulasi Silang Klasifikasi Hasil Analisis LORENS dengan *Cross Validation* untuk Konsumen *Low Price*

		Kelas Aktual			
		Threshold Optimum		Threshold 0,5	
		<i>p</i> (+)	<i>p</i> (+)	<i>p</i> (+)	<i>p</i> (+)
Kelas Prediksi	1 Partisi	<i>p</i> (+)	227717	102918	227717
		<i>n</i> (-)	65385	103980	103980
	2 Partisi	<i>p</i> (+)	227733	102952	258794
		<i>n</i> (-)	65369	103946	34308
	3 Partisi	<i>p</i> (+)	228590	105277	261325
		<i>n</i> (-)	64512	101621	31777
	4 Partisi	<i>p</i> (+)	258580	140052	278465
		<i>n</i> (-)	34522	66846	14637
	5 Partisi	<i>p</i> (+)	258969	140664	289082
		<i>n</i> (-)	34133	66234	4020

Berikut ini merupakan hasil klasifikasi analisis LORENS untuk konsumen *Medium Price*.

Tabel 4.4.4 Tabulasi Silang Klasifikasi Hasil Analisis LORENS dengan *Cross Validation* untuk Konsumen *Medium Price*

		Kelas Aktual			
		Threshold Optimum		Threshold 0,5	
		<i>p</i> (+)	<i>n</i> (-)	<i>p</i> (+)	<i>n</i> (-)
Kelas Prediksi	1 Partisi	<i>p</i> (+)	214223	42704	243882
		<i>n</i> (-)	58860	93023	29201
	2 Partisi	<i>p</i> (+)	240415	62122	252758
		<i>n</i> (-)	32668	73605	20325
	3 Partisi	<i>p</i> (+)	251161	82317	272161
		<i>n</i> (-)	21922	53410	922
	4 Partisi	<i>p</i> (+)	251781	85115	273032
		<i>n</i> (-)	21302	50612	51
	5 Partisi	<i>p</i> (+)	267068	125786	273050
		<i>n</i> (-)	6015	9941	33

Sedangkan berikut ini merupakan hasil klasifikasi untuk data *High Price* untuk masing-masing ukuran partisi.

Tabel 4.4.5 Tabulasi Silang Klasifikasi Hasil Analisis LORENS dengan *Cross Validation* untuk Konsumen *High Price*

		Kelas Aktual			
		Threshold Optimum		Threshold 0,5	
		p(+)	n(-)	p(+)	n(-)
Kelas Prediksi	1 Partisi	p(+)	297059	129204	303007
		n(-)	89372	194264	83424
	2 Partisi	p(+)	304091	144277	304693
		n(-)	82340	179191	81738
	3 Partisi	p(+)	304690	145222	304819
		n(-)	81741	178246	81612
	4 Partisi	p(+)	304728	145249	332366
		n(-)	81703	178219	54065
	5 Partisi	p(+)	304803	145676	345869
		n(-)	82340	304803	40562

Sama seperti analisis LORENS dengan *Holdout*, dari hasil klasifikasi di atas, kemudian dicari ukuran ketepatan klasifikasi yaitu akurasi, *Sensitivity* dan *Specificity*. Berikut ini merupakan nilai akurasi, *Sensitivity*, dan *Specificity* dari analisis LORENS data *Low Price*, *Medium Price*, dan *High Price* dengan metode evaluasi *Cross Validation*.

Tabel 4.4.6 Ukuran Ketepatan Klasifikasi Hasil Analisis LORENS dengan *Cross Validation*

		Threshold Optimum				
Produk	Partisi	1	2	3	4	5
Low Price	Akurasi	66,34	66,34	66,04	65,09	65,04
	Sensitivity	77,69	77,70	77,99	88,22	88,35
	Specificity	50,26	50,24	49,12	32,31	32,01
Medium Price	Akurasi	75,16	76,81	74,50	73,97	67,76
	Sensitivity	78,45	88,04	91,97	92,20	97,80
	Specificity	68,54	54,23	39,35	37,29	7,32
High Price	Akurasi	69,21	68,08	68,03	68,03	72,78
	Sensitivity	76,87	78,69	78,85	78,86	78,73
	Specificity	60,06	55,40	55,10	55,10	67,66

Layaknya analisis LORENS dengan *Holdout*, hasil dari analisis LORENS dengan *Cross Validation* pada tabel di atas memperlihatkan bahwa baik untuk produk *Low Price*, *Medium Price*, maupun *High Price*, nilai akurasi dan *specificity* mengecil

seiring semakin besarnya partisi. Sedangkan nilai *sensitivity* semakin membesar seiring besarnya ukuran partisi. *Threshold* optimum menjaga perubahan ukuran ketepatan klasifikasi menjadi lebih kecil daripada yang bisa dihasilkan dengan *threshold* 0,5. Selanjutnya kesalahan klasifikasi dihitung, berikut ini merupakan ukuran kesalahan klasifikasi analisis di atas.

Tabel 4.4.7 Ukuran Kesalahan Klasifikasi Hasil Analisis LORENS dengan *Cross Validation*

		Threshold Optimum				
Produk	Partisi	1	2	3	4	5
Low Price	<i>Error</i>	33,66	33,66	33,96	34,91	34,96
	FNR	22,31	22,30	22,01	11,78	11,65
	FPR	49,74	49,76	50,88	67,69	67,99
Medium Price	<i>Error</i>	24,84	23,19	25,50	26,03	32,24
	FNR	21,55	11,96	8,03	7,80	2,20
	FPR	31,46	45,77	60,65	62,71	92,68
High Price	<i>Error</i>	30,79	31,92	31,97	31,97	27,22
	FNR	23,13	21,31	21,15	21,14	21,27
	FPR	39,94	44,60	44,90	44,90	32,34

		Threshold 0,5				
Produk	Partisi	1	2	3	4	5
Low Price	<i>Error</i>	33.66	34.92	35.31	36.44	40.67
	FNR	22.31	11.71	10.84	4.99	1.37
	FPR	49.74	67.80	69.96	80.99	96.35
Medium Price	<i>Error</i>	22.81	26.39	33.24	33.12	33.16
	FNR	10.69	7.44	0.34	0.02	0.01
	FPR	47.18	64.51	99.43	99.73	99.85
High Price	<i>Error</i>	31.30	31.97	32.00	34.74	34.02
	FNR	21.59	21.15	21.12	13.99	10.50
	FPR	42.91	44.90	45.00	59.52	62.12

Sama seperti yang terjadi dalam analisis LORENS dengan *Holdout*, *threshold* optimum dapat melakukan klasifikasi lebih baik. Sehingga partisi optimum dicari dengan melihat hasil analisis LORENS dengan *threshold* optimum dan hasil analisis dengan *threshold* 0,5 tidak dilibatkan lagi. Berikut ini merupakan tabel yang memperlihatkan ukuran partisi optimum dari analisis LORENS dengan *Cross Validation*.

Tabel 4.4.8 Ukuran Partisi Optimum dari Analisis LORENS dengan *Cross Validation*

Produk	Partisi Optimum	Perubahan Nilai		
		Akurasi	<i>Sensitivity</i>	<i>Specificity</i>
<i>Low Price</i>	3	0,96	-10,23	16,81
<i>Medium Price</i>	4	6,21	-5,60	29,97
<i>High Price</i>	1	1,13	-1,82	4,66

Dalam peneitian ini ukuran partisi optimum dari LORENS dengan *Holdout* dan LORENS dengan *Cross Validation* adalah sama.

4.5 Pemilihan Metode Terbaik

Berdasarkan pasangan data *training* dan *testing* metode *Binary Logistic Regression* dan LORENS dengan *Holdout* adalah metode yang sebanding. Pertama-tama akan dibandingkan hasil klasifikasi LORENS dengan *Binary Logistic Regression* yang sama-sama menggunakan *Holdout*. Secara keseluruhan LORENS dengan *threshold* optimum dapat menghasilkan ketepatan klasifikasi lebih baik dari pada LORENS dengan *threshold* 0.5. Sehingga hasil analisis LORENS dengan *threshold* 0.5 tidak disertakan dalam perbandingan ketepatan klasifikasi di bawah. Berikut ini merupakan perbandingan ketepatan klasifikasi dari hasil kedua analisis tersebut.

Tabel 4.5.1 Perbandingan Ketepatan Klasifikasi

Produk	Metode	Akur.	<i>Sens.</i>	<i>Spec.</i>
<i>Low Price</i>	<i>Binary Logistic Regression</i>	66,54	78,05	50,22
	LORENS -3 Partisi	66,25	78,24	49,27
<i>Medium Price</i>	<i>Binary Logistic Regression</i>	77,32	89,59	52,61
	LORENS - 4 Partisi	74,06	92,32	37,34
<i>High Price</i>	<i>Binary Logistic Regression</i>	68,42	78,50	56,37
	LORENS – 1 Partisi	69,04	76,98	59,56

Pada data *Low Price* dan *Medium Price*, LORENS dapat meningkatkan *sensitivity* dengan cukup baik. Namun LORENS juga menurunkan *spesificity* dengan cukup banyak, sehingga akurasi dari LORENS pada kedua data tersebut tidak lebih baik dari yang dihasilkan *Binary Logistic Regression*. Meskipun begi-

tu, perbedaan akurasi dari hasil dua analisis pada dua data tersebut tidak terlalu signifikan. Sementara itu, LORENS berperan lebih baik pada data *High Price*. Meskipun LORENS menurunkan *sensitivity*, LORENS dapat meningkatkan *spesificity* dan akurasi. Namun begitu, perbedaan tingkat ketepatan pada dua hasil analisis di data *High Price* tersebut juga tidak terlihat signifikan.

Karena hasil analisis dengan *Binary Logistic Regression* diragukan, maka ketepatan klasifikasi ditentukan pada hasil analisis LORENS meskipun nilai tersebut sedikit tidak lebih memuaskan dari yang dihasilkan analisis *Binary Logistic Regression*. Cara analisis terbaik ditentukan dari kombinasi metode, jumlah partisi optimum dan *threshold* optimum yang dapat memberikan hasil ketepatan klasifikasi terbaik.

Secara keseluruhan, ketepatan klasifikasi terbaik LORENS didapatkan oleh konsumen produk *Medium Price* dengan ukuran 4 partisi dan *threshold* optimum, sedangkan ketepatan klasifikasi terendah didapatkan oleh konsumen produk *Low Price* dengan ukuran 3 partisi dan *threshold* optimum, sedangkan konsumen produk *High Price* mendapatkan ketepatan klasifikasi terbaik dengan ukuran 1 partisi dan *threshold* optimum. Di data *High Price*, semua variabel dilibatkan dalam model yang sama seperti yang dilakukan pada *Binary Logistics Regression*. Namun dengan *threshold* optimum, LORENS dapat menghasilkan ketepatan klasifikasi lebih baik daripada *Binary Logistic Regression*.

Cross Validation adalah metode yang dianjurkan oleh LORENS. Meskipun begitu, hasil analisis dengan *Cross Validation* pada data *Low Price* dan *Medium Price* tidak lebih baik dari dua hasil analisis sebelumnya. Namun pada data *High Price*, metode *Cross Validation* dapat menghasilkan ketepatan klasifikasi dengan lebih baik pada akurasi, *sensitivity*, dan *spesificity*. Metode terbaik dari penelitian ini adalah LORENS dengan *threshold* optimum dengan ukuran partisi 3 untuk konsumen *Low Price*, 4 partisi untuk konsumen *Medium Price* dan 1 partisi untuk konsumen *High Price*.

Halaman ini sengaja dikosongkan.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan pada penelitian ini, maka dapat diambil kesimpulan sebagai berikut.

1. Berdasarkan analisis *Binary Logistic Regression*, model yang menjelaskan hubungan Akumulasi Update, Harga Produk, Jawaban Kontrak, Tipe Konsumen, dan Status Pengiriman terhadap Pembelotan Konsumen masing-masing produk ialah sebagai berikut.

- a. Konsumen Produk *Low Price*

$$\pi(x) = \frac{e^{1,54-0,43X_1-2,92X_3-0,85X_5}}{1 + e^{1,54-0,43X_1-2,92X_3-0,85X_5}}$$

Konsumen yang melakukan *update* cenderung untuk tidak membelot 1,55 kali lebih besar daripada konsumen yang tidak melakukan *update*. Konsumen yang mengkonfirmasi lanjut menggunakan produk cenderung tidak membelot 18,63 kali daripada konsumen yang mengkonfirmasi berhenti menggunakan produk. Sedangkan konsumen yang berhasil menerima *e-mail* konfirmasi pembaharuan dengan tepat waktu cenderung tidak membelot 2,35 kali daripada konsumen yang pernah gagal menerima *e-mail*.

- b. Konsumen Produk *Medium Price*

$$\pi(x) = \frac{e^{2,9-0,705X_1-0,00011X_2-3,32X_3-0,0528X_4-0,201X_5}}{1 + e^{2,9-0,705X_1-0,00011X_2-3,32X_3-0,0528X_4-0,201X_5}}$$

Konsumen yang melakukan *update* memiliki kecenderungan untuk tidak membelot 2,02 kali lebih besar daripada konsumen yang tidak melakukan *update*. Konsumen yang mengkonfirmasi lanjut menggunakan suatu produk cenderung tidak membelot 25,44 kali daripada konsumen yang mengkonfirmasi berhenti menggunakan produk. Sedangkan konsumen yang berhasil menerima *e-mail* konfirmasi pembaharuan tepat waktu

cenderung tidak membelot 1,22 kali daripada konsumen yang gagal menerima *e-mail* konfirmasi pembaharuan produk. Variabel Harga Produk dan Tipe Konsumen tidak mempengaruhi kecenderungan konsumen membelot atau tidak, karena *odds ratio* bernilai 1.

c. Konsumen Produk *High Price*

$$\pi(x) = \frac{e^{1,68-0,538X_1-0,0000079X_2-2,84X_3-0,0345X_4-0,919X_5}}{1 + e^{1,68-0,538X_1-0,0000079X_2-2,84X_3-0,0345X_4-0,919X_5}}$$

Konsumen yang melakukan *update* cenderung untuk tidak membelot 1,71 kali lebih besar daripada konsumen yang tidak melakukan *update*. Konsumen yang mengkonfirmasi lanjut menggunakan produk cenderung tidak membelot 17,21 kali daripada konsumen yang mengkonfirmasi berhenti menggunakan produk. Sedangkan konsumen yang berhasil menerima *e-mail* konfirmasi pembaharuan tepat waktu cenderung tidak membelot 2,51 kali daripada konsumen yang gagal menerima *e-mail* konfirmasi pembaharuan produk. Variabel Harga Produk dan Tipe Konsumen tidak mempengaruhi kecenderungan konsumen membelot atau tidak, karena *odds ratio* bernilai 1.

Hasil ketepatan klasifikasi dari analisis *Binary Logistic Regression* adalah sebagai berikut.

Produk	Akurasi	Sensitivity	Spesificity
<i>Low Price</i>	66,54	77,31	68,42
<i>Medium Price</i>	78,05	89,6	78,5
<i>High Price</i>	50,22	52,61	56,37

- Salah satu kelemahan LORENS adalah tidak mampu menghasilkan model yang interpretatif, sehingga analisis dengan LORENS tidak menghasilkan model. Berikut ini merupakan hasil ketepatan klasifikasi dari analisis dengan LORENS pada masing-masing konsumen.

Produk	Holdout			Cross Validation		
	Akur.	Sens.	Spes.	Akur.	Sens.	Spes.
<i>Low Price</i>	66,25	78,24	49,21	66,04	77,99	49,12
<i>Medium Price</i>	74,06	92,32	37,34	73,97	92,2	37,29
<i>High Price</i>	69,04	76,98	59,56	69,21	76,87	60,06

3. Dari tiga produk, hasil klasifikasi tertinggi diraih oleh konsumen *Medium Price*, sedangkan hasil klasifikasi terendah diraih konsumen *High Price* pada metode *Binary Logistic Regression* dan *Low Price* pada metode LORENS. Berdasarkan nilai akurasi dan *specificity*, metode *Binary Logistic Regression* memberikan hasil lebih baik daripada LORENS dengan *Holdout* pada konsumen *Low Price* dan *High Price*.
4. Meskipun *Binary Logistic Regression* menghasilkan ketepatan klasifikasi lebih tinggi, namun metode ini gagal dalam menghasilkan analisis yang dapat dipercaya. Pada produk *Medium Price* dan *High Price*, variabel Harga Produk dan Tipe Konsumen teruji signifikan pada model namun *odds ratio* kedua variabel tersebut bernilai 1. Maka, metode yang dipilih untuk digunakan hasil analisisnya adalah LORENS dengan *Holdout* dan *threshold* optimum. Ukuran partisi untuk konsumen *Low Price* adalah 3 partisi, untuk konsumen *Medium Price* adalah 4 partisi, sedangkan untuk konsumen *High Price* adalah 1 partisi.

5.2 Saran

Baik analisis LORENS maupun *Binary Logistic Regression* belum dapat melakukan klasifikasi konsumen produk antivirus 'X' dengan lebih baik daripada metode-metode yang pernah digunakan sebelumnya. Sebaiknya perusahaan lebih mempertimbangkan hasil analisis di penelitian sebelumnya. Dalam pelaksanaan penelitian ini, terdapat kendala dalam menemukan ukuran partisi optimum yang harusnya dapat dilakukan oleh LORENS terlepas dari besar kecilnya ukuran prediktor terhadap banyaknya pengamatan. Hal tersebut dikarenakan *package* LORENS tidak terbuka untuk pengamatan yang mengandung ukuran pengamatan lebih banyak dari prediktornya. Sebaiknya *package* LORENS dikembangkan lagi agar dapat lebih terbuka pada berbagai kondisi *Big Data*.

Halaman ini sengaja dikosongkan.

DAFTAR PUSTAKA

- Agresti, A. (2002). *Categorical Data Analysis*. New York: John Wiley & Sons, Inc.
- Catal, C. (2012). Performance Evaluation Metrics for Software Fault Prediction Studies. *Acta Polytechnica Hungari-ca*, Vol. 9, No. 4.
- Colombus, L. Predicting Enterprise Cloud Computing Growth. *Forbes* (April 9, 2013), <http://www.forbes.com/sites/louis-columbus/2013/09/04/predicting-enterprise-cloud-computing-growth/> (Diakses 10 Juli 2014).
- Hosmer, D.W., Lemeshow, S. (2000). *Applied Logistic Regression, Second Edition*. New York: John Wiley & Sons, Inc.
- King, G., Zeng, L. (2001). Logistic Regression in Rare Events Data. *Society for Political Methodology* WV006-01
- Lee, K., Ahn, H., Moon, H., Kodell, R.L., & Chen, J.J. (2013). Multinomial Logistic Regression Ensembles. *Biopharm Stat*, 23(3), 681-94, doi: 10.1080/ 10543406. 2012.756500.
- Lim, N. (2007). Classification by Ensembles from Random Partitions using Logistic Models. In: *Applied Mathematics and Statistics*, Stony Brook University.
- Lim, N., Ahn, H., Moon, H., Chen, J. J. (2010). Classification of High Dimensional Data with Ensemble of Logistic Regression Models. *Journal of Biopharmaceutical Statistics* 20:160-171.
- Lin, M., Lucas, H.C.Jr., Shmueli, G. (2013). Too Big to Fail: Large Samples and The P-Value Problem. *INFORMS*, pp 1-12 ISSN 1526-5536.
- Martono, N. P. (2014). *Customer Lifetime Value And Defection Possibility Prediction Model Using Machine Learning*. Japan: Tokyo University of Science.
- Martono, N.P., Kanamori, K., Ohwada, H. (2014). Utilizing Customer's Purchase and Contract Renewal Details to Predict Defection in the Cloud Software Industry. *Springer International Publishing Switzerland*: PKAW 2014, LNCS

- 8863, pp. 138–149.
- Prasasti, N., Okada, M., Kanamori, K., & Ohwada, H. (2013). Customer Lifetime Value and Defection Possibility Prediction Model Using Machine Learning: An Application to a Cloud-based Software Company. *Lecture Notes in Customer Science*, 8398.
- Prasasti, N., Ohwada, H. (2014). Applicability of Machine Learning Techniques in Predicting Customer Defection. In: *International Symposium on Technology Management and Emerging Technologies* (ISTMET 2014).
- Witten, I. H., Frank, E., Hall, M. A. (2011). *Data Mining : Practical Machine Learning Tools And Techniques 3rd Edition*. Burlington: Morgan Kaufmann.

LAMPIRAN

LAMPIRAN 1 *Analisis Binary Logistic Regression* Konsumen *Low Price*

Case Processing Summary

Unweighted Cases ^a		N	Percent
Selected Cases	Included in Analysis	449999	100.0
	Missing Cases	0	.0
	Total	449999	100.0
Unselected Cases		0	.0
Total		449999	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Categorical Variables Codings

		Frequency	Parameter coding
			(1)
Status_Pengiriman	0	179948	1.000
	1	270051	.000
Tipe_Konsumen	0	428860	1.000
	1	21139	.000
Jawaban_Kontrak	0	446384	1.000
	1	3615	.000

Block 1: Method = Backward Stepwise (Wald)

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	74700.198	5	.000
	Block	74700.198	5	.000

	Model	74700.198	5	.000
Step 2 ^a	Step	-.133	1	.715
	Block	74700.064	4	.000
	Model	74700.064	4	.000
Step 3 ^a	Step	-1.949	1	.163
	Block	74698.116	3	.000
	Model	74698.116	3	.000

a. A negative Chi-squares value indicates that the Chi-squares value has decreased from the previous step.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	535687.957 ^a	.153	.206
2	535688.091 ^a	.153	.206
3	535690.039 ^a	.153	.206

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Classification Table^a

Observed	Predicted		
	Pembelotan_Konsumen		Percentage Correct
	0	1	
Step 1 Pembelotan_Konsumen 0	93589	92619	50.3
1	58952	204839	77.7
Overall Percentage			66.3
Step 2 Pembelotan_Konsumen 0	93589	92619	50.3
1	58952	204839	77.7
Overall Percentage			66.3
Step 3 Pembelotan_Konsumen 0	93589	92619	50.3
1	58952	204839	77.7
Overall Percentage			66.3

a. The cut value is .500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a Akumulasi_Update	-.436	.002	3.763E4	1	.000	.647
Harga_Produk	.000	.000	1.884	1	.170	1.000
Jawaban_Kontrak(1)	-2.925	.063	2.174E3	1	.000	.054
Tipe_Konsumen(1)	-.006	.016	.133	1	.715	.994
Status_Pengiriman(1)	-.854	.007	1.475E4	1	.000	.426
Constant	1.61	.065	6.14E2	1	.000	5.003
Step 2 ^a Akumulasi_Update	-.436	.002	3.764E4	1	.000	.647
Harga_Produk	.000	.000	1.947	1	.163	1.000
Jawaban_Kontrak(1)	-2.925	.063	2.175E3	1	.000	.054
Status_Pengiriman(1)	-.854	.007	1.475E4	1	.000	.426
Constant	1.539	.063	5.97E2	1	.000	4.660
Step 3 ^a Akumulasi_Update	-.436	.002	3.774E4	1	.000	.647
Jawaban_Kontrak(1)	-2.925	.063	2.174E3	1	.000	.054
Status_Pengiriman(1)	-.855	.007	1.479E4	1	.000	.425
Constant	1.54	.063	5.98E2	1	.000	4.665

a. Variable(s) entered on step 1: Akumulasi_Update, Harga_Produk, Jawaban_Kontrak, Tipe_Konsumen, Status_Pengiriman.

LAMPIRAN 2 *Analisis Binary Logistic Regression Konsumen Medium Price*

Case Processing Summary

Unweighted Cases ^a		N	Percent
Selected Cases	Included in Analysis	367928	100.0
	Missing Cases	0	.0
	Total	367928	100.0
Unselected Cases		0	.0
Total		367928	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Categorical Variables Codings

		Frequency	Parameter coding
			(1)
Status_Pengiriman	0	117305	1.000
	1	250623	.000
Tipe_Konsumen	0	339207	1.000
	1	28721	.000
Jawaban_Kontrak	0	367027	1.000
	1	901	.000

Block 1: Method = Enter**Omnibus Tests of Model Coefficients**

		Chi-square	df	Sig.
Step 1	Step	92114.373	5	.000
	Block	92114.373	5	.000
	Model	92114.373	5	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	375588.038 ^a	.221	.308

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Classification Table^a

Observed	Predicted		
	Pembelotan_Konsumen		Percentage Correct
	0	1	
Step 1 Pembelotan_Konsumen 0	64549	57605	52.8
1	26349	219425	89.3
Overall Percentage			77.2

a. The cut value is .500

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a Akumulasi_Update	-.705	.003	6.915E4	1	.000	.494
Harga_Produk	.000	.000	2.674E3	1	.000	1
Jawaban_Kontrak(1)	-3.236	.112	827.622	1	.000	.039
Tipe_Konsumen(1)	-.053	.015	12.048	1	.001	.948
Status_Pengiriman(1)	-.201	.009	484.136	1	.000	.817
Constant	2.9	.114	647.122	1	.000	18.17

a. Variable(s) entered on step 1: Akumulasi_Update, Harga_Produk, Jawaban_Kontrak, Tipe_Konsumen, Status_Pengiriman.

LAMPIRAN 3 *Analisis Binary Logistic Regression Konsumen High Price*

Case Processing Summary

Unweighted Cases ^a		N	Percent
Selected Cases	Included in Analysis	638908	100.0
	Missing Cases	0	.0
	Total	638908	100.0
Unselected Cases		0	.0
Total		638908	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Categorical Variables Codings

		Parameter coding
		(1)
Status_Pengiriman	0	223176 1.000
	1	415732 .000
Tipe_Konsumen	0	603023 1.000
	1	35885 .000
Jawaban_Kontrak	0	633422 1.000
	1	5486 .000

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	1.375E5	5	.000
	Block	1.375E5	5	.000
	Model	1.375E5	5	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	743181.525 ^a	.194	.259

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Classification Table^a

Observed	Predicted		
	Pembelotan_Konsumen		Percentage Correct
	0	1	
Step 1 Pembelotan_Konsumen 0	166253	124868	57.1
1	75035	272752	78.4
Overall Percentage			68.7

a. The cut value is .500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a Akumulasi_Update	-.538	.002	7.514E4	1	.000	.584
Harga_Produk	.000	.000	82.202	1	.000	1.000
Jawaban_Kontrak(1)	-2.845	.052	3.048E3	1	.000	.058
Tipe_Konsumen(1)	-.035	.012	8.088	1	.004	.966
Status_Pengiriman(1)	-.920	.006	2.213E4	1	.000	.399
Constant	1.676	.054	9.632E2	1	.000	5.344

a. Variable(s) entered on step 1: Akumulasi_Update, Harga_Produk, Jawaban_Kontrak, Tipe_Konsumen, Status_Pengiriman.

LAMPIRAN 4 Model Konsumen *Low Price* LORENS dengan *Holdout* dan *Threshold* Optimum 1Partisi

Ens.	Model Regresi Logistik Partisi 1
1	$\frac{e^{1,53-0,43X_1+0,000002X_2-2,9X_3+0,0057X_4-0,85X_5}}{1 + e^{1,53-0,43X_1+0,000002X_2-2,9X_3+0,0057X_4-0,85X_5}}$
2	$\frac{e^{1,53-0,43X_1+0,000002X_2-2,9X_3+0,0057X_4-0,85X_5}}{1 + e^{1,53-0,43X_1+0,000002X_2-2,9X_3+0,0057X_4-0,85X_5}}$
⋮	⋮
10	$\frac{e^{1,53-0,43X_1+0,000002X_2-2,9X_3+0,0057X_4-0,85X_5}}{1 + e^{1,53-0,43X_1+0,000002X_2-2,9X_3+0,0057X_4-0,85X_5}}$

LAMPIRAN 5 Model Konsumen *Low Price* LORENS dengan
Holdout dan *Threshold* Optimum 2 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2
1	$\frac{e^{0,99-0,48X_1+0,000009X_2+0,03X_4}}{1+e^{0,99-0,48X_1+0,000009X_2+0,03X_4}}$	$\frac{e^{1,06-2,72X_3-1,11X_5}}{1+e^{1,06-2,72X_3-1,11X_5}}$
2	$\frac{e^{1,52-0,43X_1+0,0000013X_2-0,86X_5}}{1+e^{1,52-0,43X_1+0,0000013X_2-0,86X_5}}$	$\frac{e^{0,37-2,82X_3-0,01X_4}}{1+e^{0,37-2,82X_3-0,01X_4}}$
3	$\frac{e^{1,01-0,49X_1+0,00001X_2-3,05X_3}}{1+e^{1,01-0,49X_1+0,00001X_2-3,05X_3}}$	$\frac{e^{1,05-0,02X_4-1,12X_5}}{1+e^{1,05-0,02X_4-1,12X_5}}$
4	$\frac{e^{1,52-0,43X_1+0,01X_4-0,86X_5}}{1+e^{1,52-0,43X_1+0,01X_4-0,86X_5}}$	$\frac{e^{0,18+0,000027X_2-2,83X_3}}{1+e^{0,18+0,000027X_2-2,83X_3}}$
5	$\frac{e^{0,99-0,48X_1+0,000009X_2+0,03X_4}}{1+e^{0,99-0,48X_1+0,000009X_2+0,03X_4}}$	$\frac{e^{1,06-2,72X_3-1,11X_5}}{1+e^{1,06-2,72X_3-1,11X_5}}$
6	$\frac{e^{1,54-0,43X_1-2,92X_3-0,85X_5}}{1+e^{1,54-0,43X_1-2,92X_3-0,85X_5}}$	$\frac{e^{0,18+0,000026X_2-0,02X_4}}{1+e^{0,18+0,000026X_2-0,02X_4}}$
7	$\frac{e^{1,01-0,49X_1+0,00001X_2-3,05X_3}}{1+e^{1,01-0,49X_1+0,00001X_2-3,05X_3}}$	$\frac{e^{1,05-0,025X_4-1,12X_5}}{1+e^{1,05-0,025X_4-1,12X_5}}$
8	$\frac{e^{1,52-0,43X_1+0,0000013X_2-0,86X_5}}{1+e^{1,52-0,43X_1+0,0000013X_2-0,86X_5}}$	$\frac{e^{0,36-2,82X_3-0,01X_4}}{1+e^{0,36-2,82X_3-0,01X_4}}$
9	$\frac{e^{0,96+0,000015X_2-2,72X_3-1,11X_5}}{1+e^{0,96+0,000015X_2-2,72X_3-1,11X_5}}$	$\frac{e^{1,05-0,48X_1+0,03X_4}}{1+e^{1,05-0,48X_1+0,03X_4}}$
10	$\frac{e^{1,07-0,49X_1-3,05X_3+0,03X_4}}{1+e^{1,07-0,49X_1-3,05X_3+0,03X_4}}$	$\frac{e^{0,96+0,000014X_2-1,11X_5}}{1+e^{0,96+0,000014X_2-1,11X_5}}$

LAMPIRAN 6 Model Konsumen *Low Price* LORENS dengan
Holdout dan *Threshold* Optimum 3 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3
1	$\frac{e^{0,18+0,000026X_2-0,02X_4}}{1+e^{0,18+0,000026X_2-0,02X_4}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{1,08-0,49X_1-3,05X_3}}{1+e^{1,08-0,49X_1-3,05X_3}}$
2	$\frac{e^{1,05-0,02X_4-1,12X_5}}{1+e^{1,05-0,02X_4-1,12X_5}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{0,18+0,000027X_2-2,83X_3}}{1+e^{0,18+0,000027X_2-2,83X_3}}$
3	$\frac{e^{0,18+0,000026X_2-0,02X_4}}{1+e^{0,18+0,000026X_2-0,02X_4}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{1,08-0,49X_1-3,05X_3}}{1+e^{1,08-0,49X_1-3,05X_3}}$
4	$\frac{e^{0,99-0,48X_1+0,0000092X_2}}{1+e^{0,99-0,48X_1+0,0000092X_2}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,36-2,82X_3-0,01X_4}}{1+e^{0,36-2,82X_3-0,01X_4}}$
5	$\frac{e^{1,52-0,43X_1-0,86X_5}}{1+e^{1,52-0,43X_1-0,86X_5}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{0,18+0,000026X_2-0,02X_4}}{1+e^{0,18+0,000026X_2-0,02X_4}}$
6	$\frac{e^{1,05-0,48X_1+0,03X_4}}{1+e^{1,05-0,48X_1+0,03X_4}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{1,06-2,71X_3-1,11X_5}}{1+e^{1,06-2,71X_3-1,11X_5}}$
7	$\frac{e^{0,36-2,82X_3-0,01X_4}}{1+e^{0,36-2,82X_3-0,01X_4}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{0,96+0,000014X_2-1,11X_5}}{1+e^{0,96+0,000014X_2-1,11X_5}}$
8	$\frac{e^{0,99-0,48X_1+0,0000092X_2}}{1+e^{0,99-0,48X_1+0,0000092X_2}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{1,06-2,71X_3-1,11X_5}}{1+e^{1,06-2,71X_3-1,11X_5}}$
9	$\frac{e^{1,08-0,49X_1-3,05X_3}}{1+e^{1,08-0,49X_1-3,05X_3}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{1,05-0,02X_4-1,12X_5}}{1+e^{1,05-0,02X_4-1,12X_5}}$
10	$\frac{e^{0,36-2,82X_3-0,01X_4}}{1+e^{0,36-2,82X_3-0,01X_4}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,99-0,48X_1+0,0000092X_2}}{1+e^{0,99-0,48X_1+0,0000092X_2}}$

LAMPIRAN 7 Model Konsumen *Low Price* LORENS dengan
Holdout dan Threshold Optimum 4Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3	Model Regresi Logistik Partisi 4
1	$\frac{e^{1,05-0,48X_1+0,03X_4}}{1+e^{1,05-0,48X_1+0,03X_4}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$
2	$\frac{e^{0,99-0,48X_1+0,0000092X_2}}{1+e^{0,99-0,48X_1+0,0000092X_2}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$
3	$\frac{e^{0,36-2,82X_3-0,015X_4}}{1+e^{0,36-2,82X_3-0,015X_4}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$
4	$\frac{e^{1,05-0,02X_4-1,12X_5}}{1+e^{1,05-0,02X_4-1,12X_5}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$
5	$\frac{e^{0,36-2,82X_3-0,015X_4}}{1+e^{0,36-2,82X_3-0,015X_4}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$
6	$\frac{e^{0,96+0,000014X_2-1,11X_5}}{1+e^{0,96+0,000014X_2-1,11X_5}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$
7	$\frac{e^{0,99-0,48X_1+0,0000092X_2}}{1+e^{0,99-0,48X_1+0,0000092X_2}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$
8	$\frac{e^{1,05-0,48X_1+0,03X_4}}{1+e^{1,05-0,48X_1+0,03X_4}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$
9	$\frac{e^{0,99-0,48X_1+0,0000092X_2}}{1+e^{0,99-0,48X_1+0,0000092X_2}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$
10	$\frac{e^{1,05-0,48X_1+0,03X_4}}{1+e^{1,05-0,48X_1+0,03X_4}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$

LAMPIRAN 8 Model Konsumen *Low Price* LORENS dengan
Holdout dan Threshold Optimum 5 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3	Model Regresi Logistik Partisi 4	Model Regresi Logistik Partisi 5
1	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$
2	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$
3	$\frac{e^{0,35-0,0077}}{1+e^{0,35-0,00}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$
4	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$
5	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$
6	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$
7	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$

8	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$
9	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$
10	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$

LAMPIRAN 9 Model Konsumen *Low Price* LORENS dengan *Holdout* dan *Threshold* 0,5 1 Partisi

Ens.	Model Regresi Logistik Partisi 1
1	$\frac{e^{1,529-0,436X_1+0,000002X_2-2,92X_3+0,0057X_4-0,85X_5}}{1+e^{1,529-0,436X_1+0,000002X_2-2,92X_3+0,0057X_4-0,85X_5}}$
2	$\frac{e^{1,529-0,436X_1+0,000002X_2-2,92X_3+0,0057X_4-0,85X_5}}{1+e^{1,529-0,436X_1+0,000002X_2-2,92X_3+0,0057X_4-0,85X_5}}$
⋮	⋮
10	$\frac{e^{1,529-0,436X_1+0,000002X_2-2,92X_3+0,0057X_4-0,85X_5}}{1+e^{1,529-0,436X_1+0,000002X_2-2,92X_3+0,0057X_4-0,85X_5}}$

LAMPIRAN 10 Model Konsumen *Low Price* LORENS dengan *Holdout* dan *Threshold* 0,5 2 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2
1	$\frac{e^{1,52-0,43X_1+0,014X_4-0,86X_5}}{1+e^{1,52-0,43X_1+0,014X_4-0,86X_5}}$	$\frac{e^{0,19+0,000027X_2-2,82X_3}}{1+e^{0,19+0,000027X_2-2,82X_3}}$
2	$\frac{e^{1,52-0,43X_1+0,014X_4-0,86X_5}}{1+e^{1,52-0,43X_1+0,014X_4-0,86X_5}}$	$\frac{e^{0,19+0,000027X_2-2,82X_3}}{1+e^{0,19+0,000027X_2-2,82X_3}}$
3	$\frac{e^{1,06-2,72X_3-0,03X_4-1,11X_5}}{1+e^{1,06-2,72X_3-0,03X_4-1,11X_5}}$	$\frac{e^{0,99-0,48X_1+0,0000092X_2}}{1+e^{0,99-0,48X_1+0,0000092X_2}}$
4	$\frac{e^{1,06-2,72X_3-0,03X_4-1,11X_5}}{1+e^{1,06-2,72X_3-0,03X_4-1,11X_5}}$	$\frac{e^{0,99-0,48X_1+0,0000092X_2}}{1+e^{0,99-0,48X_1+0,0000092X_2}}$
5	$\frac{e^{0,96+0,000014X_2-0,03X_4-1,11X_5}}{1+e^{0,96+0,000014X_2-0,03X_4-1,11X_5}}$	$\frac{e^{1,08-0,49X_1-3,05X_3}}{1+e^{1,08-0,49X_1-3,05X_3}}$
6	$\frac{e^{0,96+0,000014X_2-0,03X_4-1,11X_5}}{1+e^{0,96+0,000014X_2-0,03X_4-1,11X_5}}$	$\frac{e^{1,08-0,49X_1-3,05X_3}}{1+e^{1,08-0,49X_1-3,05X_3}}$
7	$\frac{e^{1,01-0,49X_1+0,0000098X_2-3,05X_3}}{1+e^{1,01-0,49X_1+0,0000098X_2-3,05X_3}}$	$\frac{e^{1,05-0,02X_4-1,12X_5}}{1+e^{1,05-0,02X_4-1,12X_5}}$
8	$\frac{e^{1,076-0,49X_1-3,05X_3+0,03X_4}}{1+e^{1,076-0,49X_1-3,05X_3+0,03X_4}}$	$\frac{e^{0,96+0,000014X_2-1,11X_5}}{1+e^{0,96+0,000014X_2-1,11X_5}}$
9	$\frac{e^{1,01-0,49X_1+0,0000098X_2-3,05X_3}}{1+e^{1,01-0,49X_1+0,0000098X_2-3,05X_3}}$	$\frac{e^{1,05-0,02X_4-1,12X_5}}{1+e^{1,05-0,02X_4-1,12X_5}}$
10	$\frac{e^{1,076-0,49X_1-3,05X_3+0,03X_4}}{1+e^{1,076-0,49X_1-3,05X_3+0,03X_4}}$	$\frac{e^{0,96+0,000014X_2-1,11X_5}}{1+e^{0,96+0,000014X_2-1,11X_5}}$

LAMPIRAN 11 Model Konsumen *Low Price* LORENS dengan
Holdout dan *Threshold* 0,5 3 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3
1	$\frac{e^{1,52-0,43X_1-0,86X_5}}{1+e^{1,52-0,43X_1-0,86X_5}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{0,18+0,000027X_2-2,83X_5}{e^{1,051-0,02X_4-1,12X_5}}$
2	$\frac{e^{0,99-0,48X_1+0,0000092X_2}}{1+e^{0,99-0,48X_1+0,0000092X_2}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,051-0,02X_4-1,12X_5}}{1+e^{1,051-0,02X_4-1,12X_5}}$
3	$\frac{e^{0,18+0,000026X_2-0,02X_4}}{1+e^{0,18+0,000026X_2-0,02X_4}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,52-0,43X_1-0,86X_5}}{1+e^{1,52-0,43X_1-0,86X_5}}$
4	$\frac{e^{1,051-0,02X_4-1,12X_5}}{1+e^{1,051-0,02X_4-1,12X_5}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{0,99-0,48X_1+0,0000092X_2}{e^{1,06-2,71X_3-1,11X_5}}$
5	$\frac{e^{0,18+0,000026X_2-0,02X_4}}{1+e^{0,18+0,000026X_2-0,02X_4}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{1,06-2,71X_3-1,11X_5}}{1+e^{1,06-2,71X_3-1,11X_5}}$
6	$\frac{e^{1,0768287-0,4886184X_1-3,047209X_3}}{1+e^{1,0768287-0,4886184X_1-3,047209X_3}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{0,18+0,000026X_2-0,02X_4}{e^{1,051-0,02X_4-1,12X_5}}$
7	$\frac{e^{0,36-2,82X_3-0,01X_4}}{1+e^{0,36-2,82X_3-0,01X_4}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{0,96+0,000014X_2-1,11X_5}{e^{1,051-0,02X_4-1,12X_5}}$
8	$\frac{e^{0,36-2,82X_3-0,01X_4}}{1+e^{0,36-2,82X_3-0,01X_4}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{0,96+0,000014X_2-1,11X_5}{e^{1,051-0,02X_4-1,12X_5}}$
9	$\frac{e^{0,18+0,000027X_2-2,83X_3}}{1+e^{0,18+0,000027X_2-2,83X_3}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{1,051-0,02X_4-1,12X_5}}{1+e^{1,051-0,02X_4-1,12X_5}}$
10	$\frac{e^{1,52-0,43X_1-0,86X_5}}{1+e^{1,52-0,43X_1-0,86X_5}}$	$\frac{0,18+0,000026X_2}{1+e^{0,36-2,82X_3-0,01X_4}}$	$\frac{e^{0,36-2,82X_3-0,01X_4}}{1+e^{0,36-2,82X_3-0,01X_4}}$

LAMPIRAN 12 Model Konsumen *Low Price* LORENS dengan
Holdout dan *Threshold* 0,5 4 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3	Model Regresi Logistik Partisi 4
1	$\frac{e^{1,52-0,43X_1-0,86X_5}}{1+e^{1,52-0,43X_1-0,86X_5}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{0,18+0,000026X_2-2,83X_5}{e^{1,051-0,02X_4-1,12X_5}}$
2	$\frac{e^{0,18+0,000026X_2-0,02X_4}}{1+e^{0,18+0,000026X_2-0,02X_4}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$
3	$\frac{e^{0,99-0,48X_1+0,0000092X_2}}{1+e^{0,99-0,48X_1+0,0000092X_2}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$
4	$\frac{e^{0,99-0,48X_1+0,0000092X_2}}{1+e^{0,99-0,48X_1+0,0000092X_2}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$
5	$\frac{e^{0,96+0,000014X_2-1,11X_5}}{1+e^{0,96+0,000014X_2-1,11X_5}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$
6	$\frac{e^{0,99-0,48X_1+0,0000092X_2}}{1+e^{0,99-0,48X_1+0,0000092X_2}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$
7	$\frac{e^{0,99-0,48X_1+0,0000092X_2}}{1+e^{0,99-0,48X_1+0,0000092X_2}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$

8	$\frac{e^{0,99-0,48X_1+0,0000092X_2}}{1+e^{0,99-0,48X_1+0,0000092X_2}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$
9	$\frac{e^{0,18+0,000027X_2-2,83X_3}}{1+e^{0,18+0,000027X_2-2,83X_3}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$
10	$\frac{e^{1,051-0,02X_4-1,12X_5}}{1+e^{1,051-0,02X_4-1,12X_5}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{0,18+0,000026X_2}{1+0,18+0,000026X_2}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$

LAMPIRAN 13 Model Konsumen *Low Price* LORENS dengan
Holdout dan *Threshold* 0,5 5 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3	Model Regresi Logistik Partisi 4	Model Regresi Logistik Partisi 5
1	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{0,35-0,0077X_4}}{1+e^{0,35-0,0077X_4}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$
2	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{0,35-0,0077X_4}{1+0,35-0,0077X_4}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$
3	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{0,35-0,0077X_4}{1+0,35-0,0077X_4}$
4	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{0,18+0,000026X_2}}{1+e^{0,18+0,000026X_2}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{0,35-0,0077X_4}{1+0,35-0,0077X_4}$
5	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{0,18+0,000026X_2}{1+0,18+0,000026X_2}$	$\frac{0,35-0,0077X_4}{1+0,35-0,0077X_4}$
6	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{1+0,35-0,0077X_4}{e^{1,05-1,12X_5}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{0,18+0,000026X_2}{1+0,18+0,000026X_2}$
7	$\frac{0,35-0,0077X_4}{1+0,35-0,0077X_4}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{0,18+0,000026X_2}{1+0,18+0,000026X_2}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$
8	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{1+0,18+0,000026X_2}{e^{1,05-1,12X_5}}$	$\frac{0,35-0,0077X_4}{1+0,35-0,0077X_4}$
9	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{1+0,18+0,000026X_2}{e^{1,05-1,12X_5}}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$	$\frac{0,18+0,000026X_2}{1+0,18+0,000026X_2}$	$\frac{e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$
10	$\frac{e^{0,36-2,81X_3}}{1+e^{0,36-2,81X_3}}$	$\frac{1+e^{1,05-1,12X_5}}{1+e^{1,05-1,12X_5}}$	$\frac{1+0,35-0,0077X_4}{1+0,35-0,0077X_4}$	$\frac{1+0,18+0,000026X_2}{1+0,18+0,000026X_2}$	$\frac{1+e^{1,05-0,48X_1}}{1+e^{1,05-0,48X_1}}$

LAMPIRAN 14 Model Konsumen *Medium Price* LORENS
dengan *Holdout* dan *Threshold* Optimum 1 Partisi

Ens.	Model Regresi Logistik Partisi 1
1	$\frac{e^{2,9-0,7X_1-0,00011X_2-3,24X_3-0,05X_4-0,2X_5}}{1+e^{2,9-0,7X_1-0,00011X_2-3,24X_3-0,05X_4-0,2X_5}}$
2	$\frac{e^{2,9-0,7X_1-0,00011X_2-3,24X_3-0,05X_4-0,2X_5}}{1+e^{2,9-0,7X_1-0,00011X_2-3,24X_3-0,05X_4-0,2X_5}}$
⋮	⋮
10	$\frac{e^{2,9-0,7X_1-0,00011X_2-3,24X_3-0,05X_4-0,2X_5}}{1+e^{2,9-0,7X_1-0,00011X_2-3,24X_3-0,05X_4-0,2X_5}}$

LAMPIRAN 15 Model Konsumen *Medium Price* LORENS
dengan *Holdout* dan *Threshold* Optimum 2 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2
1	$\frac{e^{2,89-0,7X_1-0,00011X_2-0,2X_5}}{1+e^{2,89-0,7X_1-0,00011X_2-0,2X_5}}$	$\frac{e^{0,7-2,78X_3+0,03X_4}}{1+e^{0,7-2,78X_3+0,03X_4}}$
2	$\frac{e^{2,26-0,69X_1-0,14X_4-0,18X_5}}{1+e^{2,26-0,69X_1-0,14X_4-0,18X_5}}$	$\frac{e^{0,89-0,000034X_2-2,77X_3}}{1+e^{0,89-0,000034X_2-2,77X_3}}$
3	$\frac{e^{1,05-2,76X_3-0,02X_4-0,49X_5}}{1+e^{1,05-2,76X_3-0,02X_4-0,49X_5}}$	$\frac{e^{0,7-2,78X_3+0,03X_4}}{1+e^{0,7-2,78X_3+0,03X_4}}$
4	$\frac{e^{2,89-0,7X_1-0,00011X_2-0,2X_5}}{1+e^{2,89-0,7X_1-0,00011X_2-0,2X_5}}$	$\frac{e^{0,7-2,78X_3+0,03X_4}}{1+e^{0,7-2,78X_3+0,03X_4}}$
5	$\frac{e^{1,05-2,76X_3-0,02X_4-0,49X_5}}{1+e^{1,05-2,76X_3-0,02X_4-0,49X_5}}$	$\frac{e^{0,7-2,78X_3+0,03X_4}}{1+e^{0,7-2,78X_3+0,03X_4}}$
6	$\frac{e^{1,05-2,76X_3-0,02X_4-0,49X_5}}{1+e^{1,05-2,76X_3-0,02X_4-0,49X_5}}$	$\frac{e^{0,7-2,78X_3+0,03X_4}}{1+e^{0,7-2,78X_3+0,03X_4}}$
7	$\frac{e^{2,74-0,71X_1-0,00011X_2-0,03X_4}}{1+e^{2,74-0,71X_1-0,00011X_2-0,03X_4}}$	$\frac{e^{1,05-2,76X_3-0,02X_4-0,49X_5}}{1+e^{1,05-2,76X_3-0,02X_4-0,49X_5}}$
8	$\frac{e^{1,05-2,76X_3-0,02X_4-0,49X_5}}{1+e^{1,05-2,76X_3-0,02X_4-0,49X_5}}$	$\frac{e^{2,74-0,71X_1-0,00011X_2}}{1+e^{2,74-0,71X_1-0,00011X_2}}$
9	$\frac{e^{2,26-0,69X_1-3,2X_3-0,17X_5}}{1+e^{2,26-0,69X_1-3,2X_3-0,17X_5}}$	$\frac{e^{0,89-0,000036X_2+0,07X_4}}{1+e^{0,89-0,000036X_2+0,07X_4}}$
10	$\frac{e^{2,74-0,71X_1-0,00011X_2-0,03X_4}}{1+e^{2,74-0,71X_1-0,00011X_2-0,03X_4}}$	$\frac{e^{1,05-2,76X_3-0,02X_4-0,49X_5}}{1+e^{1,05-2,76X_3-0,02X_4-0,49X_5}}$

LAMPIRAN 16 Model Konsumen *Medium Price* LORENS
dengan *Holdout* dan *Threshold* Optimum 3 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3
1	$\frac{e^{0,7-2,78X_3+0,03X_4}}{1+e^{0,7-2,78X_3+0,03X_4}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{1,29-0,000044X_2-0,51X_5}}{1+e^{1,29-0,000044X_2-0,51X_5}}$
2	$\frac{e^{2,74-0,71X_1-0,00011X_2}}{1+e^{2,74-0,71X_1-0,00011X_2}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{1,04-0,01X_4-0,49X_5}}{1+e^{1,04-0,01X_4-0,49X_5}}$
3	$\frac{e^{2,24-0,69X_1-0,17X_5}}{1+e^{2,24-0,69X_1-0,17X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{0,89-0,000034X_2-2,77X_3}{1+0,89-0,000034X_2-2,77X_3}$
4	$\frac{e^{1,04-0,01X_4-0,49X_5}}{1+e^{1,04-0,01X_4-0,49X_5}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{2,14-0,7X_1-3,22X_3}}{1+e^{2,14-0,7X_1-3,22X_3}}$
5	$\frac{e^{2,14-0,7X_1-3,22X_3}}{1+e^{2,14-0,7X_1-3,22X_3}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{1,04-0,01X_4-0,49X_5}}{1+e^{1,04-0,01X_4-0,49X_5}}$
6	$\frac{e^{1,05-2,76X_3-0,49X_5}}{1+e^{1,05-2,76X_3-0,49X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{2,74-0,71X_1-0,00011X_2}}{1+e^{2,74-0,71X_1-0,00011X_2}}$
7	$\frac{e^{2,14-0,7X_1-3,22X_3}}{1+e^{2,14-0,7X_1-3,22X_3}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{1,29-0,000044X_2-0,51X_5}}{1+e^{1,29-0,000044X_2-0,51X_5}}$
8	$\frac{e^{2,74-0,71X_1-0,00011X_2}}{1+e^{2,74-0,71X_1-0,00011X_2}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7-2,78X_3+0,03X_4}}{1+e^{0,7-2,78X_3+0,03X_4}}$
9	$\frac{e^{2,24-0,69X_1-0,17X_5}}{1+e^{2,24-0,69X_1-0,17X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,89-0,000034X_2-2,77X_3}}{1+e^{0,89-0,000034X_2-2,77X_3}}$

10	$\frac{e^{1,29-0,000044X_2-0,51X_5}}{1+e^{1,29-0,000044X_2-0,51X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{2,14-0,69X_1-3,22X_3}}{1+e^{2,14-0,69X_1-3,22X_3}}$
----	---	---	---

LAMPIRAN 17 Model Konsumen *Medium Price* LORENS
dengan *Holdout* dan *Threshold* Optimum 4 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3	Model Regresi Logistik Partisi 4
1	$\frac{e^{2,14-0,69X_1-0,12X_4}}{1+e^{2,14-0,69X_1-0,12X_4}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$
2	$\frac{e^{2,14-0,7X_1-3,22X_3}}{1+e^{2,14-0,7X_1-3,22X_3}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,89-0,000035X_2-0,49X_5}}{1+e^{0,89-0,000035X_2-0,49X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$
3	$\frac{e^{1,04-0,01X_4-0,49X_5}}{1+e^{1,04-0,01X_4-0,49X_5}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$
4	$\frac{e^{0,7-2,78X_3+0,03X_4}}{1+e^{0,7-2,78X_3+0,03X_4}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$
5	$\frac{e^{0,89-0,000036X_2+0,07X_4}}{1+e^{0,89-0,000036X_2+0,07X_4}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$
6	$\frac{e^{1,29-0,000044X_2-0,51X_5}}{1+e^{1,29-0,000044X_2-0,51X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$
7	$\frac{e^{2,24-0,69X_1-0,17X_5}}{1+e^{2,24-0,69X_1-0,17X_5}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$
8	$\frac{e^{2,24-0,69X_1-0,17X_5}}{1+e^{2,24-0,69X_1-0,17X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$
9	$\frac{e^{1,29-0,000044X_2-0,51X_5}}{1+e^{1,29-0,000044X_2-0,51X_5}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$
10	$\frac{e^{2,24-0,69X_1-0,17X_5}}{1+e^{2,24-0,69X_1-0,17X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$

LAMPIRAN 18 Model Konsumen *Medium Price* LORENS
dengan *Holdout* dan *Threshold* Optimum 5 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3	Model Regresi Logistik Partisi 4	Model Regresi Logistik Partisi 5
1	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$
2	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$
3	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$
4	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$
5	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$

6	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$
7	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$
8	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$
9	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$
10	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$

LAMPIRAN 19 Model Konsumen *Medium Price* LORENS dengan *Holdout* dan *Threshold* 0,5 1 Partisi

Ens.	Model Regresi Logistik Partisi 1
1	$\frac{e^{2,9-0,7X_1-0,00011X_2-3,24X_3-0,05X_4-0,2X_5}}{1+e^{2,9-0,7X_1-0,00011X_2-3,24X_3-0,05X_4-0,2X_5}}$
2	$\frac{e^{2,9-0,7X_1-0,00011X_2-3,24X_3-0,05X_4-0,2X_5}}{1+e^{2,9-0,7X_1-0,00011X_2-3,24X_3-0,05X_4-0,2X_5}}$
⋮	⋮
10	$\frac{e^{2,9-0,7X_1-0,00011X_2-3,24X_3-0,05X_4-0,2X_5}}{1+e^{2,9-0,7X_1-0,00011X_2-3,24X_3-0,05X_4-0,2X_5}}$

LAMPIRAN 20 Model Konsumen *Medium Price* LORENS dengan *Holdout* dan *Threshold* 0,5 2 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2
1	$\frac{e^{2,89-0,7X_1-0,00011X_2-0,2X_5}}{1+e^{2,89-0,7X_1-0,00011X_2-0,2X_5}}$	$\frac{e^{0,7-2,78X_3+0,03X_4}}{1+e^{0,7-2,78X_3+0,03X_4}}$
2	$\frac{e^{2,26-0,69X_1-3,2X_3-0,17X_5}}{1+e^{2,26-0,69X_1-3,2X_3-0,17X_5}}$	$\frac{e^{0,89-0,000036X_2+0,07X_4}}{1+e^{0,89-0,000036X_2+0,07X_4}}$
3	$\frac{e^{2,26-0,69X_1-3,2X_3-0,17X_5}}{1+e^{2,26-0,69X_1-3,2X_3-0,17X_5}}$	$\frac{e^{0,89-0,000034X_2-2,77X_3}}{1+e^{0,89-0,000034X_2-2,77X_3}}$
4	$\frac{e^{2,26-0,69X_1-3,2X_3-0,17X_5}}{1+e^{2,26-0,69X_1-3,2X_3-0,17X_5}}$	$\frac{e^{0,89-0,000034X_2-2,77X_3}}{1+e^{0,89-0,000034X_2-2,77X_3}}$
5	$\frac{e^{2,26-0,69X_1-3,2X_3-0,17X_5}}{1+e^{2,26-0,69X_1-3,2X_3-0,17X_5}}$	$\frac{e^{0,89-0,000034X_2-2,77X_3}}{1+e^{0,89-0,000034X_2-2,77X_3}}$
6	$\frac{e^{2,15-0,69X_1-3,22X_3-0,13X_4}}{1+e^{2,15-0,69X_1-3,22X_3-0,13X_4}}$	$\frac{1,29-0,000044X_2-0,51X_5}{1+1,29-0,000044X_2-0,51X_5}$
7	$\frac{e^{2,89-0,7X_1-0,00011X_2-0,2X_5}}{1+e^{2,89-0,7X_1-0,00011X_2-0,2X_5}}$	$\frac{e^{0,7-2,78X_3+0,03X_4}}{1+e^{0,7-2,78X_3+0,03X_4}}$
8	$\frac{e^{1,29-0,000044X_2+0,02X_4-0,51X_5}}{1+e^{1,29-0,000044X_2+0,02X_4-0,51X_5}}$	$\frac{e^{2,14-0,69X_1-3,22X_3}}{1+e^{2,14-0,69X_1-3,22X_3}}$
9	$\frac{e^{2,74-0,71X_1-0,00011X_2-0,03X_4}}{1+e^{2,74-0,71X_1-0,00011X_2-0,03X_4}}$	$\frac{e^{1,05-2,76X_3-0,49X_5}}{1+e^{1,05-2,76X_3-0,49X_5}}$

10

$$\frac{e^{2,26-0,69X_1-3,2X_3-0,17X_5}}{1+e^{2,26-0,69X_1-3,2X_3-0,17X_5}}$$

$$\frac{e^{0,89-0,000034X_2-2,77X_3}}{1+e^{0,89-0,000034X_2-2,77X_3}}$$

LAMPIRAN 21 Model Konsumen *Medium Price* LORENS
dengan *Holdout* dan *Threshold* 0,5 3 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3
1	$\frac{e^{2,14-0,69X_1-0,12X_4}}{1+e^{2,14-0,69X_1-0,12X_4}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{1,29-0,000044X_2-0,51X_5}}{1+e^{1,29-0,000044X_2-0,51X_5}}$
2	$\frac{e^{2,14-0,69X_1-0,12X_4}}{1+e^{2,14-0,69X_1-0,12X_4}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{1,29-0,000044X_2-0,51X_5}}{1+e^{1,29-0,000044X_2-0,51X_5}}$
3	$\frac{e^{1,29-0,000044X_2-0,51X_5}}{1+e^{1,29-0,000044X_2-0,51X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{2,14-0,69X_1-3,22X_3}}{1+e^{2,14-0,69X_1-3,22X_3}}$
4	$\frac{e^{2,24-0,69X_1-0,17X_5}}{1+e^{2,24-0,69X_1-0,17X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,89-0,000034X_2-2,77X_3}}{1+e^{0,89-0,000034X_2-2,77X_3}}$
5	$\frac{e^{2,74-0,71X_1-0,00011X_2}}{1+e^{2,74-0,71X_1-0,00011X_2}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{1,05-2,76X_3-0,49X_5}}{1+e^{1,05-2,76X_3-0,49X_5}}$
6	$\frac{e^{0,89-0,000036X_2+0,07X_4}}{1+e^{0,89-0,000036X_2+0,07X_4}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{1,05-2,76X_3-0,49X_5}}{1+e^{1,05-2,76X_3-0,49X_5}}$
7	$\frac{e^{1,05-2,76X_3-0,49X_5}}{1+e^{1,05-2,76X_3-0,49X_5}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,89-0,000036X_2+0,07X_4}}{1+e^{0,89-0,000036X_2+0,07X_4}}$
8	$\frac{e^{2,14-0,69X_1-3,22X_3}}{1+e^{2,14-0,69X_1-3,22X_3}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,89-0,000036X_2+0,07X_4}}{1+e^{0,89-0,000036X_2+0,07X_4}}$
9	$\frac{e^{1,05-2,76X_3-0,49X_5}}{1+e^{1,05-2,76X_3-0,49X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{2,74-0,71X_1-0,00011X_2}}{1+e^{2,74-0,71X_1-0,00011X_2}}$
10	$\frac{e^{0,89-0,000034X_2-2,77X_3}}{1+e^{0,89-0,000034X_2-2,77X_3}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{2,24-0,69X_1-0,17X_5}}{1+e^{2,24-0,69X_1-0,17X_5}}$

LAMPIRAN 22 Model Konsumen *Medium Price* LORENS
dengan *Holdout* dan *Threshold* 0,5 4 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3	Model Regresi Logistik Partisi 4
1	$\frac{e^{0,89-0,000036X_2+0,07X_4}}{1+e^{0,89-0,000036X_2+0,07X_4}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$
2	$\frac{e^{0,89-0,000036X_2+0,07X_4}}{1+e^{0,89-0,000036X_2+0,07X_4}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{2,13-0,69X_1-0,49X_5}}{1+e^{2,13-0,69X_1-0,49X_5}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$
3	$\frac{e^{2,74-0,71X_1-0,00011X_2}}{1+e^{2,74-0,71X_1-0,00011X_2}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$
4	$\frac{e^{2,74-0,71X_1-0,00011X_2}}{1+e^{2,74-0,71X_1-0,00011X_2}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$
5	$\frac{e^{1,04-0,01X_4-0,49X_5}}{1+e^{1,04-0,01X_4-0,49X_5}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$
6	$\frac{e^{1,04-0,01X_4-0,49X_5}}{1+e^{1,04-0,01X_4-0,49X_5}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$

7	$\frac{e^{0,89-0,000036X_2+0,07X_4}}{1+e^{0,89-0,000036X_2+0,07X_4}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$
8	$\frac{e^{2,74-0,71X_1-0,00011X_2}}{1+e^{2,74-0,71X_1-0,00011X_2}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$
9	$\frac{e^{0,89-0,000036X_2+0,07X_4}}{1+e^{0,89-0,000036X_2+0,07X_4}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$
10	$\frac{e^{1,29-0,000044X_2-0,51X_5}}{1+e^{1,29-0,000044X_2-0,51X_5}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$

LAMPIRAN 23 Model Konsumen *Medium Price* LORENS dengan *Holdout* dan *Threshold* 0,5 5 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3	Model Regresi Logistik Partisi 4	Model Regresi Logistik Partisi 5
1	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$
2	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$
3	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$
4	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$
5	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$
6	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$
7	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$
8	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$
9	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$
10	$\frac{e^{0,7-2,78X_3}}{1+e^{0,7-2,78X_3}}$	$\frac{e^{0,7+0,03X_4}}{1+e^{0,7+0,03X_4}}$	$\frac{e^{0,89-0,000035X_2}}{1+e^{0,89-0,000035X_2}}$	$\frac{e^{1,04-0,49X_5}}{1+e^{1,04-0,49X_5}}$	$\frac{e^{2,13-0,69X_1}}{1+e^{2,13-0,69X_1}}$

LAMPIRAN 24 Model Konsumen *High Price* LORENS dengan *Holdout* dan *Threshold* Optimum 1 Partisi

Ens.	Model Regresi Logistik Partisi 1
1	$\frac{e^{1,68-0,54X_1-0,000008X_2-2,84X_3-0,03X_4-0,92X_5}}{1+e^{1,68-0,54X_1-0,000008X_2-2,84X_3-0,03X_4-0,92X_5}}$
2	$\frac{e^{1,68-0,54X_1-0,000008X_2-2,84X_3-0,03X_4-0,92X_5}}{1+e^{1,68-0,54X_1-0,000008X_2-2,84X_3-0,03X_4-0,92X_5}}$

:	:
10	$\frac{e^{1,68-0,54X_1-0,000008X_2-2,84X_3-0,03X_4-0,92X_5}}{1+e^{1,68-0,54X_1-0,000008X_2-2,84X_3-0,03X_4-0,92X_5}}$

LAMPIRAN 25 Model Konsumen *High Price* LORENS dengan *Holdout* dan *Threshold* Optimum 2 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2
1	$\frac{e^{1,00013-0,00000046X_2-2,63X_3-1,2X_5}}{1+e^{1,00013-0,00000046X_2-2,63X_3-1,2X_5}}$	$\frac{e^{1,03-0,58X_1-0,03X_4}}{1+e^{1,03-0,58X_1-0,03X_4}}$
2	$\frac{e^{1,06-0,59X_1-2,91X_3-0,04X_4}}{1+e^{1,06-0,59X_1-2,91X_3-0,04X_4}}$	$\frac{e^{0,99-0,0000013X_2-1,2X_5}}{1+e^{0,99-0,0000013X_2-1,2X_5}}$
3	$\frac{e^{0,05+0,000016X_2-2,67X_3-0,11X_4}}{1+e^{0,05+0,000016X_2-2,67X_3-0,11X_4}}$	$\frac{e^{1,58-0,53X_1-0,92X_5}}{1+e^{1,58-0,53X_1-0,92X_5}}$
4	$\frac{e^{1,00013-0,00000046X_2-2,63X_3-1,2X_5}}{1+e^{1,00013-0,00000046X_2-2,63X_3-1,2X_5}}$	$\frac{e^{1,03-0,58X_1-0,03X_4}}{1+e^{1,03-0,58X_1-0,03X_4}}$
5	$\frac{e^{1,01-0,58X_1+0,0000025X_2-0,04X_4}}{1+e^{1,01-0,58X_1+0,0000025X_2-0,04X_4}}$	$\frac{e^{0,99-2,63X_3-1,19X_5}}{1+e^{0,99-2,63X_3-1,19X_5}}$
6	$\frac{e^{0,05+0,000016X_2-2,67X_3-0,11X_4}}{1+e^{0,05+0,000016X_2-2,67X_3-0,11X_4}}$	$\frac{e^{1,58-0,53X_1-0,92X_5}}{1+e^{1,58-0,53X_1-0,92X_5}}$
7	$\frac{e^{1,66-0,54X_1-0,0000088X_2-0,92X_5}}{1+e^{1,66-0,54X_1-0,0000088X_2-0,92X_5}}$	$\frac{e^{0,2-2,66X_3-0,1X_4}}{1+e^{0,2-2,66X_3-0,1X_4}}$
8	$\frac{e^{1,66-0,54X_1-0,0000088X_2-0,92X_5}}{1+e^{1,66-0,54X_1-0,0000088X_2-0,92X_5}}$	$\frac{e^{0,2-2,66X_3-0,1X_4}}{1+e^{0,2-2,66X_3-0,1X_4}}$
9	$\frac{e^{0,99-0,000001X_2-0,08X_4-1,2X_5}}{1+e^{0,99-0,000001X_2-0,08X_4-1,2X_5}}$	$\frac{e^{1,06-0,59X_1-2,91X_3}}{1+e^{1,06-0,59X_1-2,91X_3}}$
10	$\frac{e^{0,99-0,000001X_2-0,08X_4-1,2X_5}}{1+e^{0,99-0,000001X_2-0,08X_4-1,2X_5}}$	$\frac{e^{1,06-0,59X_1-2,91X_3}}{1+e^{1,06-0,59X_1-2,91X_3}}$

LAMPIRAN 26 Model Konsumen *High Price* LORENS dengan *Holdout* dan *Threshold* Optimum 3 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3
1	$\frac{e^{0,99-0,0000013X_2-1,2X_5}}{1+e^{0,99-0,0000013X_2-1,2X_5}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,2-2,66X_3-0,1X_4}}{1+e^{0,2-2,66X_3-0,1X_4}}$
2	$\frac{e^{1,06-0,59X_1-2,91X_3}}{1+e^{1,06-0,59X_1-2,91X_3}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,99-0,0000013X_2-1,2X_5}}{1+e^{0,99-0,0000013X_2-1,2X_5}}$
3	$\frac{e^{1,58-0,53X_1-0,92X_5}}{1+e^{1,58-0,53X_1-0,92X_5}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,2-2,66X_3-0,1X_4}}{1+e^{0,2-2,66X_3-0,1X_4}}$
4	$\frac{e^{0,05+0,000016X_2-2,66X_3}}{1+e^{0,05+0,000016X_2-2,66X_3}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{1,58-0,53X_1-0,92X_5}}{1+e^{1,58-0,53X_1-0,92X_5}}$
5	$\frac{e^{0,2-2,66X_3-0,1X_4}}{1+e^{0,2-2,66X_3-0,1X_4}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{1,01-0,58X_1+0,0000023X_2}}{1+e^{1,01-0,58X_1+0,0000023X_2}}$
6	$\frac{e^{0,2-2,66X_3-0,1X_4}}{1+e^{0,2-2,66X_3-0,1X_4}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{1,58-0,53X_1-0,92X_5}}{1+e^{1,58-0,53X_1-0,92X_5}}$

7	$\frac{e^{0,05+0,000015X_2-0,1X_4}}{1+e^{0,05+0,000015X_2-0,1X_4}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{1,06-0,59X_1-2,91X_3}}{1+e^{1,06-0,59X_1-2,91X_3}}$
8	$\frac{e^{1,0313-0,58X_1-0,03X_4}}{1+e^{1,0313-0,58X_1-0,03X_4}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,05+0,000016X_2-2,66X_3}}{1+e^{0,05+0,000016X_2-2,66X_3}}$
9	$\frac{e^{0,05+0,000015X_2-0,1X_4}}{1+e^{0,05+0,000015X_2-0,1X_4}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{1,58-0,53X_1-0,92X_5}}{1+e^{1,58-0,53X_1-0,92X_5}}$
10	$\frac{e^{1,01-0,58X_1+0,0000023X_2}}{1+e^{1,01-0,58X_1+0,0000023X_2}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,98-0,08X_4-1,2X_5}}{1+e^{0,98-0,08X_4-1,2X_5}}$

LAMPIRAN 27 Model Konsumen *High Price* LORENS dengan *Holdout* dan *Threshold* Optimum 4 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3	Model Regresi Logistik Partisi 4
1	$\frac{e^{0,05+0,000016X_2-2,66X_3}}{1+e^{0,05+0,000016X_2-2,66X_3}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{1,033-0,58X_1}}{1+e^{1,033-0,58X_1}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$
2	$\frac{e^{1,58-0,53X_1-0,92X_5}}{1+e^{1,58-0,53X_1-0,92X_5}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$
3	$\frac{e^{1,58-0,53X_1-0,92X_5}}{1+e^{1,58-0,53X_1-0,92X_5}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$
4	$\frac{e^{0,05+0,000016X_2-2,66X_3}}{1+e^{0,05+0,000016X_2-2,66X_3}}$	$\frac{e^{1,033-0,58X_1}}{1+e^{1,033-0,58X_1}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$
5	$\frac{e^{1,58-0,53X_1-0,92X_5}}{1+e^{1,58-0,53X_1-0,92X_5}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$
6	$\frac{e^{0,2-2,66X_3-0,1X_4}}{1+e^{0,2-2,66X_3-0,1X_4}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$
7	$\frac{e^{0,98-0,08X_4-1,2X_5}}{1+e^{0,98-0,08X_4-1,2X_5}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$
8	$\frac{e^{0,98-0,08X_4-1,2X_5}}{1+e^{0,98-0,08X_4-1,2X_5}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$
9	$\frac{e^{0,05+0,000016X_2-2,66X_3}}{1+e^{0,05+0,000016X_2-2,66X_3}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{1,033-0,58X_1}}{1+e^{1,033-0,58X_1}}$
10	$\frac{e^{1,01-0,58X_1+0,0000023X_2}}{1+e^{1,01-0,58X_1+0,0000023X_2}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$

LAMPIRAN 28 Model Konsumen *High Price* LORENS dengan *Holdout* dan *Threshold* Optimum 5 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3	Model Regresi Logistik Partisi 4	Model Regresi Logistik Partisi 5
1	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$
2	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$

3	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,18-0,09X_4+}}{1+e^{0,18-0,09X_4+}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$
4	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$
5	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$
6	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$
7	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$
8	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$
9	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$
10	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$

LAMPIRAN 29 Model Konsumen *High Price* LORENS dengan
Holdout dan *Threshold* 0,5 1 Partisi

Ens.	Model Regresi Logistik Partisi 1
1	$\frac{e^{1,68-0,54X_1-0,000008X_2-2,84X_3-0,03X_4-0,92X_5}}{1+e^{1,68-0,54X_1-0,000008X_2-2,84X_3-0,03X_4-0,92X_5}}$
2	$\frac{e^{1,68-0,54X_1-0,000008X_2-2,84X_3-0,03X_4-0,92X_5}}{1+e^{1,68-0,54X_1-0,000008X_2-2,84X_3-0,03X_4-0,92X_5}}$
⋮	⋮
10	$\frac{e^{1,68-0,54X_1-0,000008X_2-2,84X_3-0,03X_4-0,92X_5}}{1+e^{1,68-0,54X_1-0,000008X_2-2,84X_3-0,03X_4-0,92X_5}}$

LAMPIRAN 30 Model Konsumen *High Price* LORENS dengan
Holdout dan *Threshold* 0,5 2 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2
1	$\frac{e^{1,66-0,54X_1-0,0000088X_2-0,92X_5}}{1+e^{1,66-0,54X_1-0,0000088X_2-0,92X_5}}$	$\frac{e^{0,2-2,66X_3-0,1X_4}}{1+e^{0,2-2,66X_3-0,1X_4}}$
2	$\frac{e^{1,01-0,58X_1+0,0000025X_2-0,04X_4}}{1+e^{1,01-0,58X_1+0,0000025X_2-0,04X_4}}$	$\frac{e^{0,99-2,63X_3-1,19X_5}}{1+e^{0,99-2,63X_3-1,19X_5}}$
3	$\frac{e^{1,66-0,54X_1-0,0000088X_2-0,92X_5}}{1+e^{1,66-0,54X_1-0,0000088X_2-0,92X_5}}$	$\frac{e^{0,2-2,66X_3-0,1X_4}}{1+e^{0,2-2,66X_3-0,1X_4}}$
4	$\frac{e^{1,000668-2,63X_3-0,09X_4-1,19X_5}}{1+e^{1,000668-2,63X_3-0,09X_4-1,19X_5}}$	$\frac{e^{1,01-0,58X_1+0,0000023X_2}}{1+e^{1,01-0,58X_1+0,0000023X_2}}$
5	$\frac{e^{1,6-0,54X_1-2,84X_3-0,91X_5}}{1+e^{1,6-0,54X_1-2,84X_3-0,91X_5}}$	$\frac{e^{0,05+0,000015X_2-0,1X_4}}{1+e^{0,05+0,000015X_2-0,1X_4}}$
6	$\frac{e^{1,000668-2,63X_3-0,09X_4-1,19X_5}}{1+e^{1,000668-2,63X_3-0,09X_4-1,19X_5}}$	$\frac{e^{1,01-0,58X_1+0,0000023X_2}}{1+e^{1,01-0,58X_1+0,0000023X_2}}$

7	$\frac{e^{0,05+0,000016X_2-2,67X_3-0,11X_4}}{1+e^{0,05+0,000016X_2-2,67X_3-0,11X_4}}$	$\frac{e^{1,58-0,53X_1-0,92X_5}}{1+e^{1,58-0,53X_1-0,92X_5}}$
8	$\frac{e^{1,58-0,5X_1-0,03X_4-0,92X_5}}{1+e^{1,58-0,5X_1-0,03X_4-0,92X_5}}$	$\frac{e^{0,05+0,000016X_2-2,66X_3}}{1+e^{0,05+0,000016X_2-2,66X_3}}$
9	$\frac{e^{1,6-0,54X_1-2,84X_3-0,91X_5}}{1+e^{1,6-0,54X_1-2,84X_3-0,91X_5}}$	$\frac{e^{0,05+0,000015X_2-0,1X_4}}{1+e^{0,05+0,000015X_2-0,1X_4}}$
10	$\frac{e^{1,00013-0,00000046X_2-2,63X_3-1,2X_5}}{1+e^{1,00013-0,00000046X_2-2,63X_3-1,2X_5}}$	$\frac{e^{1,03-0,58X_1-0,03X_4}}{1+e^{1,03-0,58X_1-0,03X_4}}$

LAMPIRAN 31 Model Konsumen *High Price* LORENS dengan
Holdout dan *Threshold* 0,5 3 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3
1	$\frac{e^{1,58-0,53X_1-0,92X_5}}{1+e^{1,58-0,53X_1-0,92X_5}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,2-2,66X_3-0,1X_4}}{1+e^{0,2-2,66X_3-0,1X_4}}$
2	$\frac{e^{1,06-0,59X_1-2,91X_3}}{1+e^{1,06-0,59X_1-2,91X_3}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,98-0,08X_4-1,2X_5}}{1+e^{0,98-0,08X_4-1,2X_5}}$
3	$\frac{e^{0,99-0,0000013X_2-1,2X_5}}{1+e^{0,99-0,0000013X_2-1,2X_5}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{1,031-0,58X_1-0,03X_4}}{1+e^{1,031-0,58X_1-0,03X_4}}$
4	$\frac{e^{0,2-2,66X_3-0,1X_4}}{1+e^{0,2-2,66X_3-0,1X_4}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{1,01-0,58X_1+0,0000023X_2}}{1+e^{1,01-0,58X_1+0,0000023X_2}}$
5	$\frac{e^{0,99-2,63X_3-1,19X_5}}{1+e^{0,99-2,63X_3-1,19X_5}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{1,01-0,58X_1+0,0000023X_2}}{1+e^{1,01-0,58X_1+0,0000023X_2}}$
6	$\frac{e^{0,995-2,63X_3-1,19X_5}}{1+e^{0,995-2,63X_3-1,19X_5}}$	$\frac{e^{1,033-0,58X_1}}{1+e^{1,033-0,58X_1}}$	$\frac{e^{0,05+0,000015X_2-0,1X_4}}{1+e^{0,05+0,000015X_2-0,1X_4}}$
7	$\frac{e^{0,99-0,0000013X_2-1,2X_5}}{1+e^{0,99-0,0000013X_2-1,2X_5}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,2-2,66X_3-0,1X_4}}{1+e^{0,2-2,66X_3-0,1X_4}}$
8	$\frac{e^{1,01-0,58X_1+0,0000023X_2}}{1+e^{1,01-0,58X_1+0,0000023X_2}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,2-2,66X_3-0,1X_4}}{1+e^{0,2-2,66X_3-0,1X_4}}$
9	$\frac{e^{0,99-0,0000013X_2-1,2X_5}}{1+e^{0,99-0,0000013X_2-1,2X_5}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{1,031-0,58X_1-0,03X_4}}{1+e^{1,031-0,58X_1-0,03X_4}}$
10	$\frac{0,99-0,0000013X_2-1,2X_5}{1+0,99-0,0000013X_2-1,2X_5}$	$\frac{0,19-2,66X_3}{1+0,19-2,66X_3}$	$\frac{1,031-0,58X_1-0,03X_4}{1+1,031-0,58X_1-0,03X_4}$

LAMPIRAN 32 Model Konsumen *High Price* LORENS dengan
Holdout dan *Threshold* 0,5 4 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3	Model Regresi Logistik Partisi 4
1	$\frac{e^{1,58-0,53X_1-0,92X_5}}{1+e^{1,58-0,53X_1-0,92X_5}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$
2	$\frac{e^{1,06-0,59X_1-2,91X_3}}{1+e^{1,06-0,59X_1-2,91X_3}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$
3	$\frac{e^{1,031-0,58X_1-0,03X_4}}{1+e^{1,031-0,58X_1-0,03X_4}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$

4	$\frac{e^{1,06-0,59X_1-2,91X_3}}{1+e^{1,06-0,59X_1-2,91X_3}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$
5	$\frac{e^{1,06-0,59X_1-2,91X_3}}{1+e^{1,06-0,59X_1-2,91X_3}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$
6	$\frac{e^{0,05+0,000016X_2-2,66X_3}}{1+e^{0,05+0,000016X_2-2,66X_3}}$	$\frac{e^{1,033-0,58X_1}}{1+e^{1,033-0,58X_1}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$
7	$\frac{e^{1,01-0,58X_1+0,0000023X_2}}{1+e^{1,01-0,58X_1+0,0000023X_2}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$
8	$\frac{e^{1,06-0,59X_1-2,91X_3}}{1+e^{1,06-0,59X_1-2,91X_3}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$
9	$\frac{e^{0,05+0,000016X_2-2,66X_3}}{1+e^{0,05+0,000016X_2-2,66X_3}}$	$\frac{e^{1,033-0,58X_1}}{1+e^{1,033-0,58X_1}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$
10	$\frac{e^{0,05+0,000015X_2-0,1X_4}}{1+e^{0,05+0,000015X_2-0,1X_4}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{1,033-0,58X_1}}{1+e^{1,033-0,58X_1}}$

LAMPIRAN 33 Model Konsumen *High Price* LORENS dengan
Holdout dan Threshold 0,5 5 Partisi

Ens.	Model Regresi Logistik Partisi 1	Model Regresi Logistik Partisi 2	Model Regresi Logistik Partisi 3	Model Regresi Logistik Partisi 4	Model Regresi Logistik Partisi 5
1	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$
2	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$
3	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$
4	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$
5	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$
6	$\frac{e^{0,04+0,000015X_2}}{1+e^{0,04+0,000015X_2}}$	$\frac{e^{0,98-1,2X_5}}{1+e^{0,98-1,2X_5}}$	$\frac{e^{0,18-0,09X_4}}{1+e^{0,18-0,09X_4}}$	$\frac{e^{1,03-0,58X_1}}{1+e^{1,03-0,58X_1}}$	$\frac{e^{0,19-2,66X_3}}{1+e^{0,19-2,66X_3}}$
7	$\frac{0,04+0,000015X_2}{1+0,04+0,000015X_2}$	$\frac{0,18-0,09X_4}{1+0,18-0,09X_4}$	$\frac{0,19-2,66X_3}{1+0,19-2,66X_3}$	$\frac{0,98-1,2X_5}{1+0,98-1,2X_5}$	$\frac{1,03-0,58X_1}{1+1,03-0,58X_1}$
8	$\frac{0,19-2,66X_3}{1+0,19-2,66X_3}$	$\frac{0,04+0,000015X_2}{1+0,04+0,000015X_2}$	$\frac{1,03-0,58X_1}{1+1,03-0,58X_1}$	$\frac{0,18-0,09X_4}{1+0,18-0,09X_4}$	$\frac{0,98-1,2X_5}{1+0,98-1,2X_5}$
9	$\frac{0,18-0,09X_4}{1+0,18-0,09X_4}$	$\frac{0,19-2,66X_3}{1+0,19-2,66X_3}$	$\frac{0,04+0,000015X_2}{1+0,04+0,000015X_2}$	$\frac{0,98-1,2X_5}{1+0,98-1,2X_5}$	$\frac{1,03-0,58X_1}{1+1,03-0,58X_1}$
10	$\frac{0,98-1,2X_5}{1+0,98-1,2X_5}$	$\frac{0,19-2,66X_3}{1+0,19-2,66X_3}$	$\frac{1,03-0,58X_1}{1+1,03-0,58X_1}$	$\frac{0,04+0,000015X_2}{1+0,04+0,000015X_2}$	$\frac{0,18-0,09X_4}{1+0,18-0,09X_4}$

LAMPIRAN 34 *Threshold* Optimum dari *Cross Validation*
Konsumen *Low Price*

<i>Fold</i>	Partisi 1	Partisi 2	Partisi 3	Partisi 4	Partisi 5
1	0,5429644	0,5432767	0,5430922	0,5431867	0,5430367
2	0,5431400	0,5432311	0,5431100	0,5430589	0,5430722
3	0,5432711	0,5427633	0,5429944	0,5432511	0,5431889
4	0,5429289	0,5432344	0,5431389	0,5430600	0,5433678
5	0,5430789	0,5429400	0,5430411	0,5433133	0,5430300
6	0,5431111	0,5430300	0,5431811	0,5430678	0,5430578
7	0,5431478	0,5431656	0,5431800	0,5428578	0,5432111
8	0,5430667	0,5431344	0,5429856	0,5430000	0,5431478
9	0,5432267	0,5433289	0,5430944	0,5431644	0,5429956
10	0,5430844	0,5429156	0,5432022	0,5430600	0,5429122

LAMPIRAN 35 *Threshold* Optimum dari *Cross Validation*
Konsumen *Medium Price*

<i>Fold</i>	Partisi 1	Partisi 2	Partisi 3	Partisi 4	Partisi 5
1	0,5840454	0,5841677	0,5838497	0,5840889	0,5839897
2	0,5839489	0,5841025	0,5841392	0,5839680	0,5840359
3	0,5840876	0,5839557	0,5839748	0,5842466	0,5840414
4	0,5838158	0,5839014	0,5837166	0,5838375	0,5839680
5	0,5838660	0,5840305	0,5839870	0,5840645	0,5839924
6	0,5841297	0,5841079	0,5841011	0,5841881	0,5839951
7	0,5839816	0,5840155	0,5838728	0,5838511	0,5840128
8	0,5838973	0,5838511	0,5842819	0,5840155	0,5839924
9	0,5841365	0,5839625	0,5839517	0,5836731	0,5838389
10	0,5840658	0,5838796	0,5840998	0,5840414	0,5841079

LAMPIRAN 36 *Threshold* Optimum dari *Cross Validation*
 Konsumen *High Price*

<i>Fold</i>	Partisi 1	Partisi 2	Partisi 3	Partisi 4	Partisi 5
1	0,5220591	0,5221201	0,522027	0,5222156	0,5220943
2	0,522196	0,5221577	0,5224238	0,5222148	0,5221882
3	0,5223025	0,5220074	0,5220372	0,522232	0,5223048
4	0,5221021	0,5223385	0,5220489	0,5221115	0,5221913
5	0,5222086	0,5221342	0,522099	0,5221718	0,5221068
6	0,5220309	0,5222328	0,5223197	0,5221248	0,5222633
7	0,5223886	0,522117	0,5221859	0,5221898	0,5223745
8	0,5222	0,5220027	0,5222133	0,5222258	0,5221467
9	0,522142	0,5223651	0,5221655	0,5222352	0,522016
10	0,5221025	0,5222567	0,522212	0,5220109	0,5220461

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini, mahasiswa Jurusan Statistika FMIPA ITS:

Nama : Ayu Asfihani

NRP : 1311100084

Menyatakan bahwa data yang digunakan dalam Tugas Akhir/~~Thesis~~ ini merupakan data sekunder yang diambil dari penelitian/~~buku/Tugas Akhir/Thesis/publikasi lainnya~~ yaitu:

Sumber : Penelitian Dosen

Keterangan : Klasifikasi Pembelotan Konsumen Antivirus Norton menggunakan LORENS

Surat Pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui
Pembimbing Tugas Akhir

Surabaya, 13 Juli 2015



Dr. rer. pol. Heri Kuswanto, M.Si
NIP: 19810326 200312 1 004



Ayu Asfihani
1311100084

*(coret yang tidak perlu)

BIODATA PENULIS



Penulis, Ayu Asfihani, lahir pada tanggal 26 September 1993 di kota Gresik dari pasangan H. A. Muhdi dan Hj. Maizun. Setelah menempuh pendidikan menengah atas di SMA Negeri 1 Gresik, penulis melanjutkan studinya di jurusan Statistika Institut Teknologi Sepuluh Nopember. Surabaya.

Selain belajar di bangku kuliah, penulis juga aktif dalam kegiatan non-akademik, diantaranya ikut berkontribusi sebagai staff PST Statistika 2012/2013, staff PSDM PSM ITS 2012/2013, crew ITS EXPO, serta menjadi Wakil Ketua PSM ITS 2013/2014. Dalam hal kepelatihan, penulis telah mengikuti LKMM Pra-TD 2011, LKMM TD 2012, Pelatihan Manajemen Organisasi LMB ITS 2013. Kontribusi lain penulis dalam PSM ITS antara lain juga menjadi penyanyi aktif pada konser-konser dan perlombaan nasional maupun internasional yang diikuti PSM ITS, serta berbagai kepanitiaan lainnya. Selain aktif di dalam kampus, penulis juga telah menjadi salah satu dari peserta pertukaran pelajar pada Cultural Student Exchange oleh University of Technology of Malaysia.

Sebelum penulis menyelesaikan studinya di bangku kuliah, penulis telah memiliki berbagai pengalaman kerja, antara lain freelance di media massa Jawa Pos dan Badan Energi Nuklir Indonesia, serta menjadi intern di Perusahaan Pelabuhan Indonesia III.

Semoga Tugas Akhir ini dapat bermanfaat bagi berbagai pihak. Segala saran dan kritik yang membangun akan sangat diharapkan oleh penulis. Penulis dapat dihubungi melalui surat elektronik di ayuasfihani@gmail.com.

Halaman ini sengaja dikosongkan.